DEFENCE R&D DÉFENSE

Individual Differences:

A Review

Tonya L. Stokes-Hendriks

Defence R&D Canada
Technical Report
DCIEM TR 2001-010
March 2002

National Defence Défense nationale

Canada

Individual Differences:

A Review

Tonya L. Stokes-Hendriks

Defence and Civil Institute of Environmental Medicine

Technical Report

DCIEM TR 2001-010

March 2002

Author

Tonya L. Stokes-Hendriks

Approved by

Dr. Ross Pigeau

Section Head / Human Factors of Command Systems Section

Approved for release by

K.M. Sutton

Chair, Document Review and Library Committee

Abstract

Differential psychology refers to the objective and quantitative investigation of individual differences in behaviour. While the notion that individual differences exist can be dated back to Plato in 400 B.C., the systematic study of individual differences and psychometrics did not begin until the early 1900's. The following paper will: a) trace the history of individual differences research, b) explore the methods of measuring individual differences, c) identify the causes of individual differences, d) provide evidence attesting to the long term stability of individual differences, e) provide evidence that individual differences in intelligence and personality traits can be used to predict success in both civilian and military jobs, and f) provide evidence that individual differences in intelligence and personality traits can be used to predict leadership and creativity. Finally, this paper will highlight some of the criticisms of the current individual differences research and suggest future directions for researchers in this field.

Résumé

La psychologie différentielle renvoie à la notion d'investigation objective et quantitative des différences individuelles dans le comportement. Si l'existence du concept de différences individuelles peut être datée du temps de Plato, en 400 av. J.C., l'étude systématique des différences individuelles et de la psychométrie n'a commencé qu'au début des années 1900. Le document suivant : a) retracera l'histoire de la recherche sur les différences individuelles, b) examinera les méthodes pour mesurer les différences individuelles, c) définira les causes des différences individuelles, d) procurera une preuve attestant la stabilité à long terme des différences individuelles, e) fournira la preuve que les différences individuelles dans l'intelligence et les traits de personnalité peuvent servir à prévoir le succès dans les emplois aussi bien civils que militaires et f) fournira la preuve que les différences individuelles dans l'intelligence et les traits de personnalité peuvent servir à pronostiquer des qualités de chef et la créativité. Enfin, ce document fera ressortir certaines des critiques de la recherche menée actuellement sur les différences individuelles et proposera de futures orientations pour les chercheurs dans ce domaine.

This page intentionally left blank.

Executive summary

Differential psychology refers to the objective and quantitative investigation of individual differences in behaviour (Anastasi, 1958). While the notion that individual differences exist can be dated back to Plato in 400 B.C., the systematic study of individual differences and psychometrics did not begin until the early 1900's. The modern period of differential psychology may be viewed as beginning with the work Galton, who contributed to the field of individual differences research by advocating the measurement and quantification of both physical and mental traits on a large scale. Galton's view that a measure of intellectual functioning could be obtained through tests of sensory discrimination and reaction time (RT), however, was criticized by Binet and Henri (1895) as being too sensory and as concentrating on simple, specialized abilities. Shortly afterwards, Binet and Henri invented the first practically useful test of intelligence, which was the progenitor of all subsequent individual tests of intelligence and profoundly influenced the item contents of group tests as well.

The greatest boost for the use of tests for selection purposes came during World War I when 1.7 million American men were tested on either the Army Alpha or Army Beta for selection and classification purposes (Furnham, 1992). To date, mental tests provide a primary method for studying individual differences. They are used for placing children in schools, promotion, educational guidance, selection of students for college, and selection of workers for industry. In addition, intelligence tests may be regarded as the most important innovation in developing and legitimating the idea of individual differences, as intelligence tests were able to scientifically demonstrate the existence of such differences.

Studies that have examined the causes of individual differences have found that both genetic and environmental factors influence the development of ability and personality traits. It should be noted, however, that for some traits, individual differences appear to be primarily shaped by genetic factors. For example, Vernon's (1979) summary of the available evidence suggests that 60% of the variability in intellectual functioning can be attributed to genetic influences. Similarly, Zuckerman (1991) has found that the main personality traits of extraversion, neuroticism, and psychoticism are also primarily shaped by genetic factors. In addition, the finding that genetic components become more influential with age also attests to the overwhelming influence that genes exert on individual differences.

Results from a number of longitudinal studies attest to the long-term stability of individual differences in intellectual functioning. Although it has been shown that mean levels of intelligence test scores decline slightly with age, individuals do not change in terms of their rank ordering of test performance. Furthermore, the rank ordering of individual differences in intellectual functioning has been shown to remain stable for intervals of 45 years and more (Arbuckle, Maag, Pushkar, & Chaikelson, 1998).

Similarly, the research on stability of individual differences in personality traits suggests that trait consistency across time is uniformly high. More specifically, results from a number of sources indicate that personality traits from all five factors are stable, that stability coefficients are higher for adults over age 30 than for younger adults, and that stability gradually decreases over long retest intervals. These findings fit neatly with those found for individual differences

in intellectual functioning, as longitudinal studies on individual differences in intelligence also indicate that individual differences are stable over time, that stability coefficients are higher for adults over age 30 than for children or younger adults, and that stability gradually decreases over long retest intervals.

Because individual differences in personality characteristics and intelligence show temporal stability, personality and intelligence measures can serve as valid predictors of future behaviour. Today, personality and mental ability tests are used as selection devices to screen out unacceptable applicants and to reduce the high costs of interviewing. Hence, one of the benefits of psychological tests as a personnel selection strategy is that high selection validity translates into considerable financial savings for most organizations. This is especially important for the military, where the cost of training is extremely high. As a result, the military is one of the largest consumers of personality assessment and intelligence tests (Gal, 1995).

To date, there is considerable evidence that intellectual ability is the single best predictor of performance in virtually every job. In addition, evidence suggests that broad personality traits are also relevant to predicting success in a wide array of jobs. As a result, it is possible that a combination of cognitive ability and personality measures can predict job performance even better than the best measures of ability or personality considered alone. More research, however, is needed in this area.

Research on leadership has also found that intellectual ability is the single best predictor of leadership perceptions. Research has shown that although effective leaders are typically brighter than their subordinates are, they should not be too much brighter. Stogdill (1974) explained this finding as being due to the leader's inability to communicate effectively with his or her subordinates.

Results from studies that have attempted to determine the specific traits that clearly differentiate leaders from followers have found that intelligence, self-confidence, determination, integrity, and sociability are the main traits that consistently identify effective leaders (Northouse, 1997). As found with the personnel selection literature, more research is needed in the area.

Finally, the extent to which individual differences in intelligence and personality characteristics predict creativity is far from clear. To date, preliminary evidence suggests that creative individuals are characterized as being intelligent, independent, introverted, and flexible, as well as unstable, irresponsible, disorderly, and careless.

In sum, individual differences in intelligence and personality characteristics can be used to predict training success and performance ratings for both military and civilians. To date, individual differences in intellectual ability appears to be the single best predictor of success for both personnel selection and leadership ratings. More research is needed, however, to examine the extent to which a combination of ability and personality measures can improve our ability to predict job performance and future leaders.

Stokes-Hendriks, T.L..2000. Individual Differences. DCIEM TR 2001-010. Defence and Civil Institute of Environmental Medicine.

Sommaire

La psychologie différentielle renvoie à la notion d'investigation objective et quantitative des différences individuelles dans le comportement (Anastasi, 1958). Si la notion de l'existence des différences individuelles date pratiquement du temps de Plato, en 400 av. J.C., l'étude systématique des différences individuelles et de la psychométrie n'a vu le jour qu'au début des années 1900. On peut dire que l'époque moderne de la psychologie différentielle a commencé avec l'œuvre de Galton qui a contribué à la recherche sur les différences individuelles en proposant la mesure et la quantification à grande échelle des traits à la fois physiques et mentaux. Le point de vue de Galton selon lequel une mesure du fonctionnement intellectuel pourrait être obtenue par des tests de discrimination sensorielle et de temps de réaction a toutefois été critiqué par Binet et Henri (1895) comme étant trop sensoriel et trop axé sur de simples capacités techniques. Peu de temps après, Binet et Henri ont conçu le premier test d'intelligence pratiquement utile, qui s'est trouvé à la base de tous les tests d'intelligence individuels ultérieurs et qui a aussi fortement influencé le contenu des éléments des tests collectifs.

L'utilisation de tests à des fins de sélection a connu la plus grande poussée pendant la Première Guerre mondiale où 1,7 million d'hommes américains ont subi soit le test Army Alpha ou le test Army Beta pour les besoins de sélection et de classification (Furham 1992). Jusqu'à présent, les tests mentaux présentent une méthode primaire pour étudier les différences individuelles. Ils servent au placement des enfants dans les écoles, à la promotion, à l'orientation scolaire, à la sélection d'étudiants pour les collèges et à la sélection de travailleurs pour l'industrie. De plus, on pourrait considérer les tests d'intelligence comme l'innovation la plus importante dans le développement et la légitimation de la notion de différences individuelles, puisqu'ils pouvaient prouver scientifiquement l'existence de ces dernières.

Les études qui portaient sur les causes des différences individuelles ont révélé que non seulement les facteurs génétiques mais aussi les facteurs environnementaux influaient sur le développement des aptitudes et des traits de personnalité. Il faut toutefois noter que pour certains traits, les différences individuelles semblent être déterminées essentiellement par des facteurs génétiques. Par exemple, les éléments de preuve valides exposés brièvement par Vernon (1979) impliquent que 60 p. 100 de la variation du fonctionnement intellectuel peuvent être attribués aux influences génétiques. De même, Zucherman (1991) a trouvé que les principaux traits de personnalité comme l'extraversion, les traits névrotiques et les traits psychotiques étaient aussi déterminés principalement par les facteurs génétiques. D'autre part, on a constaté que l'influence des composants génétiques augmentait avec l'âge, ce qui prouve également l'influence dominante exercée par les gènes sur les différences individuelles.

Les résultats d'un certain nombre d'études longitudinales témoignent de la stabilité à long terme des différences individuelles dans le fonctionnement intellectuel. Bien que le niveau des résultats moyens des tests d'intelligence diminue légèrement avec l'âge, les individus ne changent pas, compte tenu du classement par grandeur des résultats de leurs tests. On a démontré en outre que le classement par grandeur des différences individuelles dans le

Stokes-Hendriks, T.L. 2000. Des Différences Individuelles: Une Sommaire. TR 2001-010. Institut De Médecine Environnmentale Pour La Défense.

fonctionnement intellectuel restait stable pendant des intervalles d'au moins 45 ans (Arbuckle et autres, 1998).

Dans le même ordre d'idées, la recherche sur la stabilité des différences individuelles de traits de personnalité laisse entendre que la stabilité des traits à travers le temps est régulièrement grande. Plus précisément, les résultats tirés de plusieurs sources indiquent que les traits de personnalité émanant des cinq facteurs sont stables, que les indices de constance sont plus élevés pour les adultes âgés de plus de 30 ans que pour les adultes plus jeunes et que la stabilité diminue graduellement pendant les longs intervalles de retests. Ces résultats correspondent bien à ceux qu'on a notés sur les différences individuelles dans le contexte du fonctionnement intellectuel, étant donné que les études longitudinales menées sur les différences individuelles concernant l'intelligence montrent aussi que les différences individuelles restent stables à la longue, que les indices de constance sont plus élevés pour les adultes de plus de 30 ans que pour les enfants ou les adultes plus jeunes et que la stabilité diminue graduellement au cours de longs intervalles de retests.

Puisque les différences individuelles dans les traits de personnalité et l'intelligence présentent une stabilité temporelle, les mesures de la personnalité et de l'intelligence peuvent servir d'indicateurs prévisionnels valides du futur comportement. Aujourd'hui, on se sert des tests d'aptitudes comme des moyens de sélection pour éliminer les candidats inadmissibles et réduire les coûts élevés des entretiens. Ainsi, un des avantages des tests psychologiques utilisés comme stratégie de sélection du personnel est que pour la plupart des organisations, la validité d'une sélection supérieure se traduit par d'énormes économies financières. Cela s'avère surtout important pour les militaires dont les coûts d'entraînement sont extrêmement élevés. Il s'ensuit que les militaires sont les plus gros utilisateurs des tests d'intelligence et d'évaluation de personnalité (Gal 1995).

Jusqu'à présent, on constate de nombreux faits portant à croire que l'aptitude intellectuelle est le meilleur indicateur prévisionnel qui soit du rendement dans pratiquement tous les emplois. Il a été prouvé en outre que les traits de personnalité généraux convenaient aussi pour prédire le succès dans une vaste gamme d'emplois. Il est par conséquent possible qu'une combinaison de mesures d'aptitudes cognitives et de mesures de personnalité puisse prévoir le rendement au travail même mieux que les meilleures mesures d'aptitudes ou de personnalité considérées séparément. On a cependant besoin d'effectuer davantage de recherche dans ce domaine.

Des recherches sur les qualités de chef révèlent aussi que l'aptitude intellectuelle est le meilleur indicateur prévisionnel des qualités de chef qui soit. Des recherches ont montré que même si généralement les chefs compétents étaient plus intelligents que leurs subordonnés, ils ne devraient pas être trop intelligents. Stogdill (1974) explique cette conclusion par l'incapacité du chef à communiquer efficacement avec ses subordonnés.

Les résultats des études visant à déterminer les traits caractéristiques qui nettement distinguent les chefs des subordonnés révèlent que l'intelligence, la confidence en soi, la détermination, l'intégrité et la sociabilité sont les principales caractéristiques qui distinguent sans faillir les chefs compétents (Northouse, 1997) La documentation sur la sélection du personnel indique que d'autres recherches sont nécessaires dans ce domaine.

Enfin, on n'est pas près de voir précisément à quel point les différences individuelles dans l'intelligence et les traits de personnalité prédisent la créativité. Jusqu'à maintenant, les indications préliminaires laissent entendre que les personnes créatives sont définies comme intelligentes, indépendantes, introverties et flexibles de même qu'instables, inconsidérées, inconduites et négligentes.

Bref, les différences individuelles dans l'intelligence et les traits de personnalité peuvent servir à prévoir le succès dans la formation et les cotes de rendement tant des civils que des militaires. Jusqu'à présent, les différences individuelles relatives à l'aptitude intellectuelle semble être le meilleur indicateur prévisionnel de succès qui soit pour non seulement la sélection du personnel mais aussi les évaluations des qualités de chef. Toutefois, des recherches sont encore nécessaires pour étudier à quel point une combinaison de mesures d'aptitudes et de personnalité peut améliorer notre capacité de prévoir le rendement au travail et repérer les futurs chefs.

This page intentionally left blank.

Table of contents

List of tables

This page intentionally left blank.

1. History of Individual Differences

Research and theories in psychology have tended to share the broad aim of understanding generalities in human behaviour (Tyler, 1995). Most branches of psychology are therefore dedicated to explaining what human beings share in common. For example, developmental psychologists study how human behaviour develops from birth to adulthood, cognitive psychologists study how humans think, process information, and solve problems, and social psychologists study the general rules which govern human social interaction (Brunas-Wagstaff, 1998). It is important to note, however, that although humans are similar in many ways, differences between individuals are even more apparent (Brunas-Wagstaff, 1998).

Differential psychology refers to the objective and quantitative investigation of individual differences in behaviour (Anastasi, 1958). Individual differences psychology is not just a single area of psychology, as it is possible to study individual differences in relation to any of the phenomena which psychologists typically study, including perception, thinking, problem-solving, and social interaction (Brunas-Wagstaff, 1998).

1.1 Early Historical Antecedents

While the scientific study of individual differences did not begin until the early 1900's, the notion that individual differences exist between human beings has been recognized throughout the centuries. The earliest instance of the explicit recognition of individual differences can be dated back to 400 years B.C. in the Republic of Plato (Anastasi, 1958). In his Republic, he both recognized differences and set up a plan for an ideal state which took these differences into consideration. The aim of Plato's ideal state was the assignment of individuals to the special tasks for which they were suited (Anastasi, 1958). According to Plato, "no two persons are born exactly alike, but each differs from each in natural endowments, one being suited for one occupation and another for another" (Book II of the Republic, 11, p.60; cited in Anastasi, 1958). Hence, Plato recognized a diversity of capacity and a diversity of accomplishment, arguing that the ideal state should be composed of such diversity (Gilliland & Clark, 1939).

During the Middle Ages, individual differences received little attention; the social context throughout the earlier centuries of European history did not favour the development of the study of the individual (Buss & Poley, 1979). In the Middle Ages, "individual" meant "inseparable"; that is, the individual was described in terms of the group to which he belonged (Williams, 1961). In the words of Williams (1961),

> The complexity of the term [individual] is at once apparent in this history, for it is the unit that is being defined, yet defined in terms of its membership of a class. The separable entity is being defined by a word that has meant 'inseparable'...The crucial history of the modern description is a change in emphasis which enables us to think of 'the individual' as a kind of absolute

> without immediate reference...to the group of which he is a member (p. 73; cited in Buss & Poley, 1979).

In contrasting this era with contemporary life, Campbell (1956) wrote:

> Then all meaning was in the group, in the great anonymous forms, none in the self-expressive individual; today no meaning is in the group-none in the world: all is in the individual (p.338; cited in Buss & Poley, 1979).

1.2 The Social Construction of Individual Differences

Buss and Poley (1979) argued that by the nineteenth century, the forces of democracy and capitalism in Britain paved the way for the study of individual differences through the creation of specialized occupations. The capitalistic society created individual differences in the sense of providing opportunities for previously unheard of specialization of human talent. Capitalism at this time had produced mainly two large classes – the bourgeoisie and the proletariat (i.e., the working class and the middle class). Within each of these separate classes, there was a high degree of occupational diversity, with individual differences serving as a necessary prerequisite for the development of the modern differentiated state. Marks (1976-77) argued that industrial needs created an industrial class structure that was rationalized by individual differences; the fact of individual differences did not necessitate a particular class structure.

In addition, Buss and Poley (1979) argued that the measurement and quantification of individual differences may be viewed in the context of reflecting capitalistic values, in which measurement and quantification play such an important role in determining salaries, prices, losses, profits, markets, and the like. They stated that just as it is possible to measure and quantify man's products, it is possible to measure and quantify man himself. Hence, the rise of capitalistic economy produced a new image of man, which paved the way for the measurement and quantification of psychological characteristics.

1.3 First Systematic Measurement of Individual Differences

As noted by Anastasi (1958), the first systematic measurement of individual differences was undertaken in 1816 by an astronomer named Bessel. Bessel was interested in an incident that had occurred in 1796, whereby Maskelyne, an astronomer at the Greenwich observatory, fired his assistant, Kinnebrook, for observing times of stellar transits almost a second later than he did. As a result, Bessel developed the 'personal equation', which referred to the difference in seconds between estimates of two observers, and became the first researcher to record quantitative measurements of individual differences.

1.4 Rise of Experimental Psychology

A great, yet indirect, influence on the problem of individual differences arose from the work of Fechner and Weber in the mid 1800's (Gilliland & Clark, 1939). Although they were interested in the relation between mind and body, they developed methods of psychophysical

measurement that have since been adapted for the measurement of individual differences (Gilliland & Clark, 1939). More specifically, they developed the method of just noticeable differences (a.k.a., method of limits) which allows one to determine how much two stimuli must vary to be distinguished as different, and the method of constant stimuli, whereby each variable is compared several times with a standard and the point at which 50% of the judgments are correct is considered the threshold (Gilliland & Clark, 1939). While Fechner and Weber never used these methods to determine individual differences, many researchers have since adopted these methods to make such comparisons. For example, Thurstone (1928) adopted some of Fechner's techniques in measuring attitudes.

In 1879, Wundt established the first experimental psychology laboratory in Leipzig where he studied visual and auditory sensation, reaction time (RT), and psychophysics (Anastasi, 1958). Wundt and other early experimental psychologists sought to discover general laws that would apply to all humans, ignoring individual differences or regarding them as chance errors (Tyler, 1995). Early experimental psychologists frowned upon individual differences, as the more individual variation, the less accurate the generalizations regarding its nature would be (Brunas-Wagstaff, 1998). Hence, the time between the Middle Ages and the middle of the 19th century was characterized by general principles that made no allowance for individual differences.

The systematic study of individual differences and psychometrics did not begin until the early 1900's. The modern period of differential psychology may be viewed as beginning with the work of Sir Francis Galton (1822-1911). Galton contributed to the field of individual differences research by advocating the measurement and quantification of both physical and mental traits on a large scale (Buss & Poley, 1979). Furthermore, one of Galton's greatest contributions to the study of individual differences was the development of statistical procedures for describing variation in traits among individuals; of specific significance is his advancement of the correlation coefficient and the concept of regression toward the mean (Buss & Poley, 1979).

Galton also wrote extensively on the explanation or interpretation of why individuals differ. Galton was particularly interested in individual differences in intelligence and he put much effort in his attempt to document the hereditary basis for such differences (Buss & Poley, 1979). For Galton, it became apparent that to measure and analyze individual differences, it was first necessary to invent mental tests to tap the various mental functions; his tests, however, largely tapped sensory functions rather than the more cognitive functions usually associated with intelligence (Buss & Poley, 1979).

Galton's belief that one could estimate the level of intellectual functioning through the measurement of sensory capacities (i.e., discrimination) is exemplified in the following quote,

> the only information that reaches us concerning outward events appears to pass through the avenue of our senses; and the more perceptive the senses are of difference, the larger is the field upon which our judgment and intelligence can act (1884, p.27; cited in Buss & Poley, 1979).

Accordingly, Galton believed that 'the sensory discriminative capacity would be highest among the intellectually ablest' (1884, p.29; cited in Buss & Poley, 1979). Thus, the testing movement began with the pioneer research of Galton and his use of simple sensory tests.

Another contributor to the development of the psychological testing movement was Cattell (1860-1904), who contributed to both the rise of experimental psychology and the measurement of individual. Cattell's first elaborate exploration of individual differences was in the use of freshman tests conducted at Columbia in 1894 (Murphy, 1938). These tests, which were administered annually to college students in an attempt to measure their intellectual level, included measures of muscular strength, keenness of vision and hearing, sensitivity to pain, RT, memory, and the like. Hence, Cattell shared Galton's view that a measure of intellectual functioning could be obtained through tests of sensory discrimination and RT (Anastasi, 1958).

Shortly thereafter, in 1895, Binet and Henri (1895) criticized early tests as being too sensory and as concentrating on simple, specialized abilities (see Anastasi, 1958, for a review). Instead, they proposed to measure intelligence more directly by using more complex tests measuring memory, imagery, imagination, attention, comprehension, strength of will, and motor skills.

In 1904, the French Minister of Public Instruction appointed a commission to study the problem of retardation among public school children (Anastasi, 1958). Thus, Binet and Simon's initial efforts were directed toward the problem of identifying those children whose intellectual capacity was such that they could not profit from the traditional Parisian school system (Buss & Poley, 1979). As a result, Binet and Simon (1905) prepared the first intelligence scale designed to yield an overall index of an individual's level of intellectual functioning. This test (after revisions by Binet in 1908 and 1911) received a major revision by Terman at Stanford University in 1916 to produce the Stanford-Binet Intelligence Test, which is still in common use today. Such tests attracted attention from psychologists all over the world, and as a result, the tests were translated into many languages (Anastasi, 1958).

In North America, however, the development of the mental testing movement received its impetus from the outbreak of World War I (WWI). In 1917, Robert M. Yerkes, the president of the American Psychological Association, proposed that psychologists could be of service to the military, stating that:

> In the present perilous situation it is obviously desirable that the psychologists...act unitedly in the interests of our defense. Our knowledge and our methods are of importance to the military service of our country, and it is our duty to cooperate to the fullest extent and immediately toward the increased efficiency of our Army and Navy (1921, p.7; cited in Marks, 1976-77).

The rapid classification of the 1.7 million new recruits for the war effort was urgently required if the diverse abilities and talents were to be maximally used (Buss & Poley, 1979). The Binet-Simon tests, also known as individual tests, were inadequate to serve this purpose because only one subject could be tested at a time. Consequently, a number of professional psychologists were trained in military testing and designed suitable group intelligence tests for the military (Marks, 1976-77). As a result, literate soldiers were tested with the Army Alpha and illiterate or foreign soldiers were tested with the Army Beta.

The advent of group intelligence scales, with their ease of administration and scoring, was a primary factor in the popularization of psychological testing (Anastasi, 1958). Furthermore, the use of the army tests brought psychology worldwide acceptance. Not only were the tests the first major practical achievement of psychology used on a wide scale, but they also helped to make psychology respectable, since it could now claim to be scientific and quantifiable (Marks, 1976-77).

Thus, psychology's involvement with the military, and hence, the advent of group intelligence tests, served to legitimize psychology as a science. Psychology's success was illustrated in the rapid growth of the American Psychological Association from 31 members in 1892 to 442 members in 1922, 60% of whom were involved in work related to mental testing (Marks, 1976-77).

After WWI, there was a dramatic increase in the number of available tests, with group intelligence scales being developed for all ages and types of subjects (Anastasi, 1958). To date, mental tests provide a primary method for studying individual differences. They are used for placing children in schools, promotion, educational guidance, selection of students for college, and selection of workers for industry (Gilliland & Clark, 1939).

In short, intelligence tests may be regarded as the most important innovation in developing and legitimizing the idea of individual differences (Marks, 1976-77). Intelligence tests "scientifically" demonstrated the existence of individual differences and provided the rational for numerous programs based on such differences (Marks, 1976-77).

This page intentionally left blank.

2. Measuring Individual Differences

Psychometrics is the branch of psychology concerned with the measurement of individual differences. The psychometric approach to intelligence and personality follows from the assumption made by psycho-physicists, such as Weber, that is possible to measure aspects of human psychological functioning (Brunas-Wagstaff, 1998). The principle behind psychometric testing is that it is possible to quantify implicit characteristics of people through scores on ability tests or self-reports of their feelings, thoughts, and behaviour (Brunas-Wagstaff, 1998). The tool used by psychometric psychologists to measure ability or personality is the psychometric test. Accordingly, psychological tests are often used as operational definitions of abstract psychological concepts (Cooper, 1998).

Cooper (1998) noted that it is difficult or impossible to devise items that are pure measures of a trait, since individual's responses to a single test item are influenced by a whole host of traits, states, attitudes, moods, and the like. He argues that the total score on the test will inevitably be a better estimate of an individual's trait score because if each item is affected by a different set of nuisance factors, the influence of the nuisance factors will tend to cancel out, allowing the influence of the trait to accumulate.

2.1 Reliability

To explore the dependability and accuracy of a trait measure, reliability analysis is typically performed. Reliability analysis indicates the range of fluctuation in one's score as a result of irrelevant chance factors (Dobreva-Martinova, 1999). Reliability of a scale is estimated by calculating a reliability coefficient. Cronbach's alpha, a measure of the reliability or internal consistency of a test, indicates the extent to which the items on a particular test measure a given trait (Howell, 1992). To understand reliability, one must first be familiar with the concept of correlation. The correlation coefficient (r) refers to the strength of the relation between two variables. Correlation coefficients range between –1.00 and +1.00. A high positive correlation indicates that as values in one variable increase, the values in the other variable also increase. Conversely, a high negative correlation indicates that as values in one variable increase, the values in the other variable decrease. A correlation of zero indicates that no relation exists between the variables.

Cronbach's alpha is influenced by the average correlation between pairs of items on a test. If all of the items in a test measure the same trait, then the correlation between pairs of items will be large and positive, producing a high measure of reliability (Howell, 1992). Accordingly, one widely applied rule of thumb is that a test should not be used if it has a reliability coefficient below 0.7, and that it should not be used for important decisions about an individual (e.g., for assessment of the need for remedial education) if it has a reliability coefficient less than 0.9 (Cooper, 1998). Furthermore, the higher the value of alpha, the less error there is in the measurement of the trait (Cooper, 1998).

Reliability is an important issue for individual differences researchers who conduct correlational analyses. For example, suppose we obtained a correlation of 0.5 between

inspection time (IT), a measure of speed of information processing, and intelligence. In addition, suppose the reliability coefficients for the IT paradigm and the intelligence test are .65 and .70, respectively. Although significantly greater than zero, should we be concerned that our obtained correlation between IT and intelligence is much smaller than 1.00, the maximum possible correlation? The answer is no because the reliability of the two tests will limit the maximum correlation expected between the two sets of scores (Cooper, 1998). More specifically, the largest correlation that one can expect between two tests is the product of the square roots of their reliabilities (Cooper, 1998). Hence, in our example, the maximum correlation that one can expect to get between these two tasks is .67.

2.2 Validity

A test is said to be valid if it measures what it claims to measure (see Bernstein & Nunnally, 1994 for a review of validity). Strictly speaking, one validates the use to which a measuring instrument is put rather than the instrument itself; hence, tests can be valid for one purpose but not for another (Bernstein & Nunnally, 1994). For example, a test used to select first year university students may be valid for that purpose, yet be invalid for use in graduate school admissions.

Validity has been given three major meanings: (1) content validity, (2) predictive validity, and (3) construct validity. Content validity is the least empirical of the three approaches and depends on the extent to which authorities agree on how well the test material was sampled (Bernstein & Nunnally, 1994). For example, when devising a math test for Grade 3 students, one must ensure that a representative collection of items is used. That is, a test will not have content validity if the items of the test do not represent what was learned in the course. Face validity, which refers to the extent to which a test appears valid, differs from content validity because it is assessed after constructing the measuring instrument, whereas the latter follows from a test plan developed before the items are generated (Bernstein & Nunnally, 1994).

Construct validity deals with the measurement of psychological attributes that are abstract and latent rather than concrete and obvious, such as anxiety (Bernstein & Nunnally, 1994). In general, measures that presumably define a given construct should all intercorrelate highly.

Predictive validity can be evaluated simply in terms of how well the measure correlates with its designated criterion (Bernstein & Nunnally, 1994). It is important to note, however, that prediction can be impaired under a number of circumstances, such as when: (1) a long period elapses between obtaining the predictor and obtaining the criterion, which allows more opportunity for spurious influences to affect the outcome; (2) a predictor is ill defined or is a composite of a number of attributes, and (3) range restrictions exist on either the predictor, the criterion, or both (Bernstein & Nunnally, 1994).

2.3 Factor Analysis

Another statistical tool that lies at the heart of individual differences research is factor analysis. The method of factor analysis was first devised by Spearman (1904; cited in Brunas-Wagstaff, 1998) to establish whether it is meaningful to distinguish between different forms of intellectual ability. More specifically, Spearman wanted to know whether verbal

ability, spatial skills, numerical skills, and so on are independent of each other or whether they all share something in common. Spearman found high correlations between these variables and concluded that tests for verbal, numerical, and spatial skills all depended on a single underlying factor that he named general intelligence (g).

Factor analysis assumes that when variables are inter-correlated, they share something abstract or hypothetical in common (Brunas-Wagstaff, 1998). Thus, factor analysis enables researchers to determine if the items on a given test measure the same construct (e.g., 20 items that claim to measure anxiety) or whether they measure several different constructs (Cooper, 1998). Similarly, factor analysis can determine whether a number of different tests that claim to measure the same construct actually produce one factor, or whether several factors are required.

Cooper (1998) noted that factor analysis has three main uses in psychology. First, factor analysis may be used to construct tests. For example, 50 test items may be written to measure some ability or personality trait. By submitting subjects' responses to these items to a factor analysis, one may find that some items have high loadings on one factor while other items have high loadings on another factor. The items with strong loadings on one factor will measure the same psychological construct and thus form one scale, whereas the items that have strong loadings on another factor will form another scale. As a result, factor analysis can show how many scales run through the test and which items belong to each scale.

Second, factor analysis can be used in data reduction. That is, one may have five different tests that are believed to measure different aspects of personality. By factor analyzing these tests, one may find that these tests actually measure the same underlying psychological trait. This will allow researchers to think in terms of a single theoretical construct as compared to five theoretical concepts.

Third, factor analysis may be used for checking the psychometric properties of questionnaires. For example, before one can compare test results from two different populations or cultures, it must be shown that the same psychometric properties exist between the two groups. If the same psychometric structure is not found for both groups, then comparisons based on overall performance cannot be made between the two groups.

In short, psychometrics is the branch of psychology concerned with the measurement of individual differences. One should not attempt to make inferences about an individual based on his or her test score until both reliability and validity analyses have been conducted on the test in question. Tests that are unreliable or that lack validity produce test scores that are difficult, if not impossible, to interpret as unreliable tests produce scores that are highly influenced by chance factors, while tests that lack validity do not measure what they claim to measure.

Before leaving the topic of measurement of individual differences, it is also important to recognize that scores on mental tests do not represent amounts of anything. Instead, what they show is how the person tested compares with other people with regard to the trait under consideration (Anastasi, 1958). For example, an IQ of 50 does not represent 50 units of anything, nor does it tell us that someone with an IQ of 50 is twice as intelligent as someone with an IQ of 25 (Anastasi, 1958). The only conclusion that one can draw about this score is

that the person in question is considerably below average as compared to the population providing the norms for the test (Anastasi, 1958).

3. Causes of Individual Differences

Why do individuals differ from one another? The question of whether personality and ability are socially determined, or whether they are genetically determined, is generally seen as being one of the most important issues in psychology (Cooper, 1998). The problem of the causation of individual differences has practical significance in many fields. Any activity designed to improve behaviour development must be based upon an understanding of the factors that influence such development (Anastasi, 1958).

The major schools of personality theory have traditionally been strongly environmentalist in their explanation of individual differences (Buss & Poley, 1979). For example, Watson's (1930) extreme position on the importance of environment is summarized in his famous saying:

> Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in, and I'll guarantee to take any one at random and train him to be any kind of specialist I might select ...doctor, lawyer, artist, merchant-chief, and yes even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors (p.104).

As illustrated by this statement, most early behaviourists belittled and often denied the effects of genetic influences on behaviour.

Conversely, many early biologists and geneticists decried the effects of environmental influences on behaviour. Biological theorists suggested that all children are not equal at birth because of differences in their genetic makeup. They stressed the importance of genetic factors and biological mechanisms in determining behaviour, suggesting that such mechanisms create the propensity for certain types of behaviours to be manifested, given the appropriate environmental conditions (Cooper, 1998). For example, Galton (1869), the first psychologist to make a serious experimental attack upon the problem of the relative influence of heredity and environment, believed that most differences are due to heredity and, in support of his view, he produced a large body of evidence demonstrating that scientific eminence tends to run in families.

Today, few scientists are proponents of the extreme environmentalist or extreme hereditarian explanation of individual differences in traits, characteristics, or abilities. Instead, modern biologists, psychologists, and geneticists agree that the basis of individual differences is to be found in each individual's hereditary background and in the environmental conditions under which the individual has developed. For example, Jensen (1979) stated that an "environment-free intelligence is as inconceivable as a nutrition-free weight" (p.182).

The most widely accepted view of the heredity-environment relationship is the interaction view. Proponents of this view acknowledge that certain gene patterns may be fostered by one set of environmental conditions and other gene patterns by different sets of conditions. For example, it is possible that bright children benefit from different kinds of stimulation than do dull children, not merely from more extensive or complex stimulation (Jensen, 1969).

This approach to individual differences, known as the behaviour-genetic approach, attempts to analyze the variance in overt behaviour into the proportionate contributions of genetic variance sources, environmental variance sources, and the variance due to an interaction between genes and environment (Buss & Poley, 1979). Adherents to this view argue that it is not meaningful to ask how much of a particular individual score is due to heredity and how much to environment, as both of these influences are 100% important or required in terms of causing a given behaviour to occur (Buss & Poley, 1979). That is, one cannot have a behaviour occurring without either a biological organism to emit it or an environment in which it is to occur. On the other hand, it is meaningful to ask how much of the variability among persons on a certain variable, such as an ability, is due to environment, how much is due to heredity, and how much is to the joint effects of genotype and environment, including both their covariance and any interaction (Vernon, 1979).

3.1 Research Designs That Measure Heredity and Environment

3.1.1 Twin Studies

3.1.1.1 Monozygotic twins reared together versus reared apart

Monozygotic or identical twins develop from the division of a single fertilized ovum and thus have identical sets of genes. Identical twins share identical genotypes. Genotype refers to the sum of all the intrinsic, genetic information an individual has (Rosenzweig, Leiman, & Breedlove, 1996). On the other hand, phenotype refers to the sum of all physical characteristics that make up an individual. An individual's phenotype is determined by the interaction of its genotype and extrinsic factors, including experience (Rosenzweig et al., 1996). Thus, twins with identical genotypes do not have identical phenotypes because they have not received identical extrinsic influences.

If genetic influences are responsible for the development of a given trait, then the correlation between twins reared apart should be close to unity (Buss & Poley, 1979). The extent to which the coefficient falls below 1.0 can plausibly be attributed to differences in their pre- and postnatal conditions, their upbringing, or other environmental influences (Jensen, 1979).

3.1.1.2 Monozygotic versus dizygotic twins

Dizygotic or fraternal twins develop from two separate ovums that are fertilized at the same time by two separate sperm cells. Although born at the same time, fraternal twins are no more alike in heredity than ordinary siblings. On genetic grounds, the correlation between fraternal twins should be no different than that observed between two ordinary siblings (Jensen, 1979).

Studies of monozygotic versus dizygotic twins, whereby each set of twins is reared together, assume that monozygotic twin differences are a result of environmental influences (since monozygotic twins are genetically identical), whereas differences between dizygotic twins arise from both genetic and environmental differences (Buss & Poley, 1979). Such studies allow researchers to estimate the heritability index, which refers to the proportion of total phenotypic differences attributable to genetic influences. The heritability index is calculated by subtracting the observed variance in monozygotic twins from the observed variance in dizygotic twins (i.e., variability due to heredity) divided by the observed variance in dizygotic twins (i.e., variability due to both heredity and environment; Buss & Poley, 1979).

3.1.2 Foster-Child Studies

These studies allow one to see the extent to which the environment can influence scores on a trait or ability and consists of comparing the correlations of the child with both the biological parent and the adoptive or foster parent. On genetic grounds, the correlation between unrelated children or between foster parent and adopted children should be zero, and any departures from this prediction can generally be attributed to environmental influences (Jensen, 1979). Furthermore, if the correlation between the child and the biological parent is greater than that between the child and the foster parent, then it seems reasonable to conclude that heredity plays a strong role in accounting for individual differences variance (Buss & Poley, 1979). Hence, the more the adopted child resembles their foster parent or foster siblings on a given trait, the greater the effects of the environment, whereas the more the adopted child resembles their biological parent or biological siblings on such traits, the greater the effects of the heredity (Gilliland & Clark, 1939).

3.1.3 Experimentally Controlled Studies

These experiments, which are not conducted on humans for obvious ethical reasons, are carried out in animal studies which allow heredity or environment to be experimentally controlled. For example, Tryon (1940; cited in Anastasi, 1958) used such a design to illustrate the influence of heredity on maze running ability among white rats. In his study, he ran a group of 142 unselected rats in a maze for 19 runs. He then selected the most extremely maze-bright rats and bred them together, and the most extremely maze-dull rats and bred them together. Environment was held constant such that all of the living conditions were the same for the bright and dull groups. He repeated this procedure for eight generations. Results indicated that while variability in maze running ability between the two groups showed little differences among the first generation of bright and dull rats, the two groups became completely differentiated in maze running ability by the sixth generation. Thus, Tryon was able to provide evidence supporting the genetic influence of maze-running ability among white rats.

Crow, Neel, and Stern (1967) argue that "animal experiments have shown that almost any trait can be changed by selection", and they conclude that human intelligence could analogously be raised or lowered, though probably rather slowly and over many generations.

3.2 Causes of Intelligence: Empirical Findings

Conrad and Jones (1940; cited in Anastasi, 1958) conducted the first extensive study on the correlation between intelligence test scores of parents and children. In this study, intelligence tests were administered to 269 family groups, including 977 persons between the ages of 3 and 60. The younger subjects were tested with the Stanford-Binet intelligence test, while the older subjects were tested with the Army Alpha. The correlation found between parents and children was identical to that found between siblings, namely, 0.49. Conrad and Jones concluded that the obtained correlation of 0.49 is consistent with a hereditary interpretation of both parent-child and sibling resemblance.

Plomin's (1988) summary of the existing twin studies also found that the correlation between IQ's for identical twins reared together is 0.87, whereas the correlation for fraternal twins reared together is 0.53. The higher correlation obtained between identical twins reared together is attributed to their greater genetic similarity, thus further supporting the notion that individual differences in intelligence are, in part, genetically determined.

Vernon (1979) summarized some 52 investigations of familial correlations in measured intellectual functions collected by Jarvick and Erlenmeyer-Kimling (1967), and also noted that there is a striking similarity between the genetically expected and obtained correlations (see Table 1). He concluded that while all of the correlations support the conclusion that heredity plays a major part, environmental differences between pairs of twins or similarities in the environment of foster children (or foster child with parent) also appear to have a significant effect on intelligence. For example, Jensen's review indicated that identical twins reared together correlate 0.9 whereas identical twins reared apart correlate 0.75.

Further evidence for the influence of environmental factors on individual differences in intelligence stems from Jarvick and Erlingmeyer-Kimling's (1967) finding that the correlation between unrelated children reared together and between foster-parent and child is approximately 0.2. Similarly, Plomin (1988) reported that the correlation between IQ's of unrelated children living together and the correlation between IQ's of adoptive parents with adoptive children are 0.23 and 0.20, respectively. Such correlations, expected on genetic grounds, should be zero (Buss & Poley, 1979).

In his review of a number of adoption studies, Vernon (1979) found that studies examining the correlation between child and true-parent ability yield an approximate median coefficient of over 0.30, whereas foster-child and foster-parent ability yields a median coefficient of less than 0.20. This evidence supports the notion that both genetic and environmental influences are responsible for individual differences in intelligence.

Hence, the finding that the adopted child-true parent correlation is greater than zero indicates that individual differences in intelligence are, in part, genetically determined. On the other

hand, support that one's environment contributes to individual differences in intellectual functioning stems from the finding that foster-children and foster-parents, who share no genetic similarities, have IQ's that are positively, albeit moderately, correlated.

More recently, Plomin (1988) summarized a large amount of data from twin, family, and adoption studies and found that the correlation between IQs of pairs of identical twins reared apart is 0.74. The finding that the correlation for identical twins reared apart is less than that obtained for identical twins reared together supports the notion that environmental influences are, in part, responsible for individual differences in intelligence.

Vernon (1979) summarized a number of studies that have used the Jarvick and Erlingmeyer-Kimling (1967) data to determine the relative contribution of genetic and environmental factors to phenotypic variance. He found that such analyses lead to heritability coefficients ranging from 0.65 to 0.75. More recently, Bouchard and McGue (1981) reviewed over 140 studies on the relative importance of genetic and environmental influences contributing to individual differences in intelligence and concluded that approximately 50% of the variation in intelligence is attributable to genetic influences. Vernon's summary of the available data indicates that there is some convergence from these studies suggesting that phenotypic variance is comprised of roughly 60% genetic influence, 30% environmental influence, and 10% influence from genetic-environmental covariance.

In short, while the literature tends to indicate that IQ correlations between different familial pairings are strikingly close to that expected on genetic grounds, there is evidence to indicate than environmental influences are also responsible for individual differences in intelligence. In addition, Vernon (1979) has cited evidence from several studies to indicate that with age, the influence of genes increases whereas that of family environment decreases.

3.3 Causes of Personality: Empirical Findings

The question of whether personality traits are socially determined or whether they are substantially influenced by our genes has not been studied as extensively as intelligence. Recent research into the genetic basis of personality, however, suggests that most measurable aspects of normal personality appear to be at least moderately heritable (McGuffin & Thapar, 1992).

For decades, progress in research on trait psychology was hindered by the lack of a generally accepted model of trait structure, as different researchers examined different traits, or the same traits under different labels (Costa, Yang, & McCrae, 1998). In the 1980's, however, several different groups of researchers independently concluded that almost all traits could be understood in terms of five broad trait domains or factors, commonly labeled Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness (Costa & McCrae, 1992). These traits are collectively referred to as the "Big Five" or the Five Factor Model (FFM) of Costa and McCrae (1992).

As described by Salgado (1998), Neuroticism refers to the degree to which the individual is insecure, anxious, depressed, and emotional versus calm, self-confident, and cool. Extraversion is the extent to which individuals are gregarious, assertive, and sociable versus

reserved, timid, and quiet. Openness to Experience defines individuals who are creative, curious, and cultured versus practical with narrow interests. Agreeableness refers to the degree to which individuals are cooperative, warm, and agreeable versus cold, disagreeable, and antagonistic. Conscientiousness is the extent to which individuals are hardworking, organized, dependable, and persevering versus lazy, disorganized, and unreliable. To date, most of the research on the FFM of personality has used the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992), a 240-item questionnaire with 30 scales that measure six different facets of each of the five factors.

Zuckerman (1991) summarized a number of older studies and found that the genetic influence on the personality traits of extraversion/sociability, neuroticism/emotionality, and psychoticism/impulsivity was substantial. For example, pairs of identical twins had very similar levels of extraversion (*r* values ranging from 0.46 to 0.61), whereas pairs of fraternal twins were much less alike on this trait (*r* values ranging from 0.06 to 0.42). Similarly, he found that pairs of identical twins had very similar levels of neuroticism (*r* values ranging from 0.33 to 0.54) and psychoticism (*r* values ranging from 0.47 to 0.70), whereas pairs of fraternal twins were much less alike on these traits (*r* values ranging from 0.07 to 0.41 and from 0.25 to 0.41 for neuroticism and psychoticism, respectively).

Using the results of Zuckerman's (1991) summary, Cooper (1998) estimated the heritability of extraversion to be in the order of 0.54 to 0.64, whereas the influence of shared environment (i.e., common family experiences) was found to be virtually zero. He noted that some personality traits have been found to be substantially influenced by common environments. For example, in their study of prosocial and antisocial behaviour, Gojone and Stevenson (1997) found that antisocial behaviour was found to have a small genetic component (0.24 vs. 0.54 for prosocial and antisocial behaviours, respectively) and was the only behaviour for which the influence of the shared environment was substantial (0.2 vs. 0.54).

Loehlin (1992) used twin and adoption studies to determine the extent to which neuroticism, extroversion, openness to experience, agreeableness, and conscientiousness, (i.e., the "Big Five" personality traits of Costa and McCrae, 1992) are influenced by genetic make-up. The heritability values for these scales ranged from about 0.3 to 0.5.

Miles and Carey (1997) performed a meta-analysis on data from 24 twin and adoption studies that used various personality measures of aggression to examine the extent to which genetic influences are responsible for individual differences in aggression. Results indicated a very strong overall genetic effect accounting for up to 50% of the variance in aggression. Self-report and parental ratings indicated that both genes and the family environment were important determinants of individual differences in aggression among youths. With age, however, the influence of genes increased whereas that of family environment decreased.

Cadoret, Cain, and Crowe (1983) used data from three adoption studies to test the extent to which a gene-environment interaction affected the development of antisocial behaviours. The first study consisted of 364 adoptees selected from psychopathological backgrounds who were separated at birth from their biological parents and relatives. Subjects in the second study were 40 adoptees who were children of incarcerated female offenders who were separated from their biological parents within their first month of life. Subjects in the third study were 108 adoptees whose biological parents had psychopathological disorders. Results showed that

subjects with both genetic and adverse environmental factors showed significant increases in antisocial behaviours. Furthermore, the increase in the number of antisocial behaviours due to both genetic and environmental factors acting together was greater than the predicted increase from either factor acting alone.

Thus, to date, several studies exist supporting the influence of both genetic and environmental influences on personality development. While some behaviours are influenced to a greater extent by environmental factors, recent evidence indicates that the three main personality traits of extraversion, neuroticism, and psychoticism are primarily shaped by genetic factors. Furthermore, the previously cited study by Cadoret et al. (1983) illustrates that an interaction between genes and environment may also influence the development of personality traits or behaviours.

Accordingly, in answer to the question "Why do individuals differ from one another", the evidence indicates that both environmental and genetic factors are responsible for variability in both personality and ability. However, evidence from both lines of research supports the notion that genetic components become more influential with age.

This page intentionally left blank.

4. Stability of Individual Differences

Humans are usually assumed to be susceptible to influence in their early years, but are thought to become increasingly stable in important respects with age, remaining resistant to change throughout most of the adult life course (Alwin, 1994). The idea that individuals are predictable beings, capable of a high degree of consistency of behaviour across situations, is a deeply held belief in most cultures (Alwin, 1994).

Accordingly, two aspects of stability can be examined: absolute and relative stability. The former refers to whether there are normative shifts in the mean levels of intelligence or personality (Costa & McCrae, 1998). Among intelligence researchers, one issue of central interest is whether intelligence declines with age. Similarly, among personality researchers, one issue of central interest is whether personality changes over time. To address these questions, the most convenient way to estimate mean level changes with age is by conducting a cross-sectional study in which individuals of different ages are compared at the same time (Costa, Yang, & McCrae, 1998).

The latter concerns the preservation of rank order of individual differences in some trait or characteristic (Costa & McCrae, 1997). Hence, the presence of stability indicates that individuals who perform well at one time perform well again at a later time, and individuals who perform poorly at first perform poorly again later. Relative stability is addressed by conducting a longitudinal study that follows the same individuals over time. It is measured as a stability coefficient, which consists of a retest correlation based on the readministration of a measure after a period of years, and is independent of the absolute level of the trait, and thus, of any developmental changes that uniformly affect all the subjects (Costa & McCrae, 1998). When the retest interval is short (i.e., a week or month), the retest correlation is interpreted as a measure of the reliability of the instrument, however, when the retest interval is long (i.e., a year or decade), the same retest correlation is usually interpreted as evidence of the stability of the trait itself (Costa, Yang, & McCrae, 1998).

It is important to note, however, that the observed retest correlation will be affected by test unreliability as well as by any real changes in the underlying trait (Costa, Yang, & McCrae, 1998; see section II of the current paper for more on reliability). Roberts and DelVecchio (2000) argue that time is the most obvious factor that undermines stability as time may degrade trait consistency because of the cumulative effects of unreliability and/or because of the experience of true change.

As this paper is focused on individual differences per se, the section on stability will focus primarily upon relative stability (i.e., the preservation of rank order of individual differences).

4.1 Stability of Intelligence: Empirical Findings

The study of stability in intellectual functioning has been motivated by both theoretical and practical concerns. The theoretical construct of intelligence (i.e., Spearman's g) was founded on the assumption that the trait being measured was a stable one (Feldman & Adams, 1989).

As a result, for a trait measure to be taken seriously, it had to be able to capture the essential features of the trait it represented not only at a particular moment in time, but over longer periods of time as well (Feldman & Adams, 1989). As a result, it was necessary to establish stability through repeated measurements in order to demonstrate that the ordinal relation in IQ was maintained through the life cycle (Feldman & Adams, 1989).

4.1.1 Instability of Infant Intelligence

One problem facing those who have attempted to establish the stability of IQ has been the measurement of intelligence in infancy. For example, Bayley (1949; cited in Bornstein, 1989) conducted a longitudinal study which involved correlating children's performance as infants, as measured on an infant test essentially derived from Binet's original mental development scale, with their performance on a standardized intelligence test at 18 years of age. Her results indicated no predictive correlation from infancy to maturity. Cronbach (1970; cited in Vernon, 1979) reported that tests given around 2 to 3 years of age correlate 0.6 to 0.7 with Stanford-Binet test scores 1 year later, but only 0.3 to 0.4 with the same test 12 years later.

Today, many investigators use the preferential looking paradigm rather than traditional infant tests to estimate infant's information processing abilities. The preferential looking method takes advantage of the fact that babies like to look at things that are interesting rather than boring; for memory, as well as for information processing in general, what's interesting is what's new. As a result, when a baby is given two stimuli to look at, a previously presented stimulus and a new stimulus, a preference for the new stimulus indicates that the baby must remember the old stimulus. Habituation refers to the decreased tendency to look at the old stimulus (i.e., preferential looking at the new stimulus).

Investigators that have used the preferential looking paradigm rather than traditional infant tests to estimate infant's information processing abilities have found moderate correlations with later performance on mental tests. For example, Bornstein (1989) reported that across 10 reports, the median predictive correlation for measures of habituation with intelligence later in child development was 0.49, suggesting that infant information processing does posses a moderate degree of predictive validity with respect to childhood cognitive functioning.

Greater stability is found, however, when children are first tested at a later age. For example, Cronbach (1970; cited in Vernon, 1979) reported Stanford-Binet test-retest correlations as high as 0.7 for children who were first tested at 4 years of age and retested 12 years later, and 0.9 for children who were first tested at 11 years of age and retested 12 years later.

Vernon (1979) suggests that there are two reasons for the weak validity of most infant tests. First, the low predictive value of infant tests may arise from the variability of performance from day to day. Vernon argues that infants are apt to be highly distractible, and much depends on their general level of activity or passivity and on the social reactions of the infant to the tester. Evidence supporting this view stems from Bayley (1949) who reported an average correlation of 0.57 between infant tests

at 3-month intervals compared to an average correlation of 0.92 at 3 year intervals when the same group of children was tested at a later age.

Second, the low predictive value of infant tests may be due to the fact that the test items selected as representative of children's developmental level in the early years are different in content from those given from about 5 years of age on. For example, tests given to infants and young children are largely psychomotor, including gross and fine movements, reactions to objects, imitation, and the beginnings of speech, whereas tests given to children 5 years of age and on are mainly based on verbal reasoning (Vernon, 1979). Evidence in support of this view stems from Hofstaetter (1954; cited in Vernon, 1979) who factor analyzed the IQs of children included in the California Growth Studies from 2 to 13 years of age. Results indicated three major factors: the first factor was identified as sensorimotor and was prominent in the tests given in infancy; the second factor was identified as persistence or rigidity and was prominent among tests given from 20 to 40 months; the third factor represented 'g' and was prominent among tests given from 4 years of age on.

It is important to note, however, that although infant test scores are typically not predictive of later performance, low scores tend to be more diagnostic than high ones since they often result from prematurity, which puts infants at an elevated risk for adverse developmental outcomes, including impaired cognition (Bacharach & Baumeister, 1998), or brain damage (Hunt, 1976; cited in Vernon, 1979). Gifted children who obtained Binet IQs of 140 or more were no different from average children on the Bayley Infant Scale at 8 months of age (McCall, 1976; cited in Vernon, 1979).

4.1.2 Stability of Mean Levels of Adult Intelligence

Considerable evidence exists that many cognitive and intellective variables tend to stabilize in early adulthood, whereas changes due to differential aging set in only very late in the adult years (Alwin, 1994). For example, Schaie's (1983) 21-year longitudinal study of aging found that while intellectual abilities were very stable prior to the age of 60, a reliable decrement had occurred for all abilities by the age of 74.

Similarly, Schaie and Hertzog's (1983) 14-year longitudinal study on intellective ability also reported negligible changes prior to age 50. In their study, subjects were tested on the Thurstone's Primary Mental Abilities Test both 7 and 14 years after it's initial administration. While the declines that were observed after 50 were small, there was an increased decline after mean age 60 on all PMA subtests, although the decline began later for the PMA subtest Verbal Meaning.

In short, most studies on developmental changes in intellective abilities suggest that intellectual functioning is stable throughout adulthood, with deterioration due to cognitive skills beginning sometime during the 60's or 70's.

4.1.3 Stability of Individual Differences in Adult Intelligence

To date, most longitudinal studies of individual differences in intellectual abilities suggest that the rank ordering of intellectual abilities remains stable over time. For example, Schaie and Strother (1968) readministered the Science Research Associates Primary Mental Abilities Test to a group 142 males and 160 females seven years after initial testing. Subjects ranged in age from 20 to 70 and were placed in the appropriate five-year age interval from 20 to 70 years of age. Although significant cross-sectional age changes were found, analysis of the comparative age gradients suggested that age changes over time within a given individual appeared to be much smaller than differences between cohorts. An uncorrected test-retest correlation of 0.94 was found over the seven-year interval.

Bradway and Thompson (1962) reported evidence indicating that individual differences in intellectual abilities remain stable over 25 years. In their study, 111 subjects who had been tested with the Stanford-Binet at the preschool level and again 10 years later at the adolescent level were administered the Stanford-Binet and Weschler Adult Intelligence Scale (WAIS) 25 years after initial testing. Correlations of preschool IQ with adult Stanford-Binet and full WAIS IQs were 0.59 and 0.64, respectively. The degree of correspondence over the 25-year period was not significantly different from the correlation of 0.65 found between preschool and adolescent Stanford-Binet IQ after an interval of 10 years. Furthermore, neither the correlation of 0.80, found between adolescent Stanford-Binet and adult WAIS, nor the correlation of 0.85, found between adolescent and adult administrations of the Stanford-Binet, differed significantly from the correlation of 0.83 that was found between Stanford-Binet and WAIS administered to the subjects on the same day when they were adults.

Schwartzman, Gold, Andres, Arbuckle, and Chaikelson (1987) administered the Revised Examination "M" Test of intelligence to 260 male Canadian army veterans approximately 40 years after they had received the test as World War II (WWII) army recruits. The M Test was developed in both French and English versions as a speeded group-administered test by the Canadian Army to aid in classifying and assigning army recruits; it has been shown to correlate highly with a number of well-standardized intelligence tests including the American Army Alpha Test ($r = 0.80$; Schwartzman et al., 1987). A test-retest reliability coefficient of 0.78 was found over the 40-year interval, attesting to the long-term stability of individual differences in intellectual functioning.

Arbuckle, Maag, Pushkar and Chaikelson (1998) also found evidence for the stability of individual differences in intellectual functioning among WWII male army veterans. Their sample consisted of 132 subjects, all of whom had M Test data from three occasions: the initial screening done by the army between 1940 and 1945, an interview done by their research team approximately 40 years later (around 1985), and a reinterview done by their team approximately 45 years after the initial screening (around 1990). Rather than examining total score test-retest stability, as done by Scwartzman et al. (1987), they examined the stability coefficients for each of the three verbal subtests (Arithmetic, Vocabulary, and Verbal Analogies) and for each of

the three nonverbal subtests (Picture Completion, Picture Anomalies, and Paper Formboard). The 45-year correlations ranged from moderate to high, with test-retest correlations of 0.73 and 0.70 found over the 40 and 45 year intervals for the verbal section of the exam, and 0.46 and 0.51 found for the nonverbal subtests. Thus, their results indicate that the stability of intellectual functioning extends over 45 years.

Owens (1966) found evidence for the stability of individual differences in Army Alpha scores over a 40-year period. His sample consisted of 96 subjects, all of whom had test scores from three occasions: initial scores collected during an entrance test at Iowa State College in 1919, retest scores collected 31 years later in 1950, and retest scores collected again 42 years after the initial test (around 1961). The uncorrected test-retest correlations between original test scores and retest scores were 0.79 and 0.77 for the 31 and 42-year intervals, respectively. In addition, results indicated that the interval between 1950 and 1962 was one of relative constancy in mental ability, as exemplified by the test-retest correlation of 0.92.

Another method of analyzing stability is by investigating the degree to which the factor structure of a test or battery of tests changes at each longitudinal measurement occasion. Using this method, the stability of individual differences across longitudinal occasions is reflected in the covariances of factors with themselves over time (Hertzog & Schaie, 1983). If the covariance of a factor at Time 1 with itself at Time 2 is large and positive, then individuals are preserving their relative order about the factor means between Times 1 and 2 (Hertzog & Schaie, 1983). On the other hand, a zero or near zero covariance would reflect a high degree of flux in individual differences between Times 1 and 2 (Hertzog & Schaie, 1983).

Hertzog and Schaie (1983) used this reasoning to examine the stability of g, as measured by Thurstone's Primary Mental Abilities battery, over a 14-year period among 412 subjects from Schaie's Seattle Longitudinal Study. Results indicated high factor covariances, which, when standardized, reflected correlations of greater than 0.9 between g at each longitudinal session. As aforementioned, high factor covariances reflect a high degree of stability in individual differences. As a result, their data suggest that there is not much change in the relative ordering of individuals on g over a 14-year period.

In short, results from a number of longitudinal studies attest to the long-term stability of individual differences in intellectual functioning. It is important to note that the fact that mean levels of intelligence test scores decline slightly with age does not necessarily imply that individuals will change in terms of their rank ordering of test performance (Costa & McCrae, 1994). Hence, the available evidence suggests that the maintenance of rank order in individual differences in intellectual functioning remains stable for intervals of 45 years and more.

4.2 Stability of Personality: Empirical Findings

Most research that has been conducted on the stability of personality has examined the stability of personality traits, that is, the dimensions of individual differences in tendencies to

show consistent patterns of thoughts, feelings, and actions (Costa & McCrae, 1997). Some have claimed that trait measures should not be used to study personality change because traits are, by definition, enduring dispositions. With such an interpretation, empirical evidence of stability is often viewed as vaguely tautological (Costa & McCrae, 1997). Costa and McCrae (1997) argue that because traits must persist across situations and over a period of days or weeks, to be distinguished from transient moods or states, does not necessarily imply that they will show long-term stability. That is, although traits must demonstrate short-term test-retest reliability, over time, a reliable measure can show either stability or change.

4.2.1 Stability of Mean Levels in Adults

Results of a number of cross-sectional studies that have examined mean level differences in the FFM of personality have consistently found that aging results in small declines in Neuroticism, Extraversion, and Openness, and small increases in Agreeableness and Conscientiousness (Costa & McCrae, 1994). Results from a national sample of nearly 10,000 men and women aged 35 to 84 found correlations ranging from –0.12 to –0.19 between age and measures of Neuroticism, Extraversion, and Openness, further supporting the view that these three traits show small declines with age (Costa & McCrae, 1994). Similarly, a recent study that tested 1539 men and women aged 21 to 64 on the NEO-PI-R reported that the correlations of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness with age were –0.12, -0.12, -0.12, +0.17, and +0.09, respectively (Costa & McCrae, 1994). Costa and McCrae (1994) note that although all of these correlations were statistically significant, age accounted for less than 3% of the variance in any of the scales.

4.2.2 Stability of Individual Differences in Personality in Adults

As found in the intelligence literature, the results of studies on stability in individual differences in personality are large and striking. For example, Simon and Thomas (1983) readministered the Eysenck Personality Inventory (EPI), which assesses two of the major dimensions of personality, Neuroticism and Extraversion, to 1087 female and 851 male students one year after its initial administration. Results showed that females scored significantly higher than males on Neuroticism measures, whereas the opposite was found for Extraversion. Stability coefficients for all groups were positive and substantial and retest coefficients of 0.79 and 0.75 were found for Neuroticism and Extraversion.

Costa and McCrae (1977) conducted a 9-year longitudinal study to examine the stability of Neuroticism and Extraversion, as measured by the EPI. Correlations between original test scores and retest scores obtained 9 years later were 0.65 and 0.76 for Neuroticism and Extraversion, suggesting that these personality traits remain stable for at least 9 years

Costa, McCrae, and Arenberg (1980) examined the stability of personality by readministering the 10 scales of the Guilford-Zimmerman Temperament Survey to a subset of 200 male participants in the Baltimore Longitudinal Study both 6 and 12 years after initial testing. The correlations between original test scores and retest

scores ranged from 0.70 to 0.86 ($M = 0.76$) and from 0.61 to 0.87 ($M = 0.73$) for the 6 and 12-year intervals, respectively. Furthermore, when stability was separated from reliability, the values ranged from 0.80 to 1.00, indicating extreme stability in personality over 12 years.

Leon, Gillum, Gillum, and Gouze (1979) periodically tested a group of middle aged men on the Minnesota Multiphasic Personality Inventory (MMPI) over a 30 year period from middle age ($M = 49$) to old age ($M = 77$). Results indicated test-retest correlations of 0.74, 0.81, and 0.59 for measures of Neuroticism, Extraversion, and Psychoticism, respectively, after a 6-year interval, whereas correlations of 0.50, 0.82, and 0.49 were found after a 13-year interval, and correlations of 0.35, 0.74, and 0.33 were found after a 30-year interval. These results indicate considerable stability of personality traits over a 30-year period.

Conley (1984) reviewed 29 longitudinal studies of the rank-order consistency of personality traits to test the relation between consistency and time and to determine whether specific traits, such as extraversion, neuroticism, or psychoticism, varied in stability. His findings indicated that personality traits were more consistent over shorter time intervals. For example, when corrected for test unreliability, stability measures of extraversion were very high, averaging 0.98 over a 1-year period, 0.70 over a 10 year period, and 0.50 over a 40 year period. The consistency of neuroticism and psychoticism measures were lower than that obtained for extraversion. Conley attributed the difference in stability coefficients among the various traits, however, to differential scale reliability, as extraversion scale measures were more internally consistent than either neuroticism or psychoticism.

Similarly, Schuerger, Zarrella, and Hotz's (1989) comprehensive review of the rank-order consistency of traits also found that longer test-retest intervals resulted in lower rank-order consistency. In addition, their findings indicated that nonclinical samples (i.e., subjects not suffering from psychopathology) were more consistent than clinical samples, and that men and women did not differ in rank-order consistency.

4.2.3 Stability of Individual Differences in Personality in Adolescence

Thus far, evidence has been presented attesting to the stability of individual differences in personality traits among adults. Research on stability of individual differences in personality traits among younger populations, however, is rather modest. For example, Carmichael and McGue found evidence for modest individual differences stability when subjects were tested on the EPI in late adolescence ($M = 16$ years of age) and 19 years later in adulthood ($M = 35$ years of age). Results indicated retest correlations of 0.35 and 0.41 for Extraversion and Neuroticism. Although significantly greater than zero, these stability coefficients are weaker than those found in longitudinal studies of individual differences in personality among adults.

Similarly, Finn's (1986) comparison of young (ages 17-25) and middle aged (ages 43-53) adults on the MMPI over a 30-year interval showed substantially higher stability for the group of adults who were over 30 during the initial testing phase. Results

indicated average test-retest stability coefficients of 0.53 for the older adults compared to 0.38 for the younger adults. While this difference was not significant, it does support the finding that the stability of individual differences in personality traits among adults over the age of 30 is higher than that found among adolescence.

Guerrin and Gottfried (1994) conducted a 10-year longitudinal study to examine the stability of temperament dimensions from infancy to preadolescence. In this study, 109 children were tested every 6 months from 1 year of age through 3.5 years and at yearly intervals from ages 5 through 12 on a parental report scale which assess temperament on nine dimensions (activity, rythmicity, approach, adaptability, intensity, mood, persistence, distractibility, and threshold). Cross-time correlations showed a pattern of statistically significant correlations across preschool years, middle childhood years, and preadolescence. For example, average correlations of 0.36, 0.41, and 0.62 were found between ages 2 and 5, 5 and 8, and 8 and 12 over the nine temperament dimensions. This finding parallels that found in the adult literature that as the age at Time 1 increases, stability coefficients increase.

Finally, Roberts and DelVecchio (2000) compiled test-retest correlation coefficients from 152 longitudinal studies. Their meta-analysis indicated that trait consistency increases in a linear yet steplike fashion from infancy to middle age, where it then reaches its peak sometime after age 50. More specifically, they state that beyond the earliest years of life, trait consistency increases in a step-like function with increases coming in the preschool years, in young adulthood, and then again in middle age. After middle age, trait consistency reaches a plateau around 0.75.

4.2.4 Criticisms of Stability of Personality Estimates

Some researchers have argued that one reason for high stability coefficients may be that individuals may wish to present a consistent image of themselves and thus intentionally recall how they answered previously, thereby inflating the stability estimates (Costa & McCrae, 1997). Woodruff (1983) examined this hypothesis by asking middle-aged men and women to recall how they answered the items on the California Test of Personality 25 years earlier. Results indicated that subjects' recollections were poor as indicated by the finding that previous scores correlated more strongly with current scores than with recalled previous scores.

Further evidence against this argument stems from longitudinal studies that have used unbiased observers to complete the personality measure on an individual at Time 2 rather than the individual himself. For example, Costa and McCrae (1988) reported 6 year test-retest correlations of 0.83, 0.77, and 0.80, respectively, for spouse ratings on the neuroticism, extraversion, and openness subtests of the NEO-PI-R. In a later study, Costa and McCrae (1992) found 7-year test-retest correlations for single peer raters ranging from 0.74 to 0.78 for men and from 0.63 to 0.84 for women. These data also support the position that personality is stable after age 30, and that the stability is not simply a result of one attempting to present a consistent image of oneself.

Another criticism often made against stability research is that stability measures of 0.65 or less indicate that Time 1 measures typically account for less than half of the variance in Time 2 measures, and thus indicate that as much change as stability occurs over time in personality traits. Costa and McCrae (1994) refute this argument, however, by noting that such an argument assumes that personality measures are error-free, which is not the case. As a result, they show that when estimates of true scores are obtained by disattenuating correlations for test-retest unreliability, stability coefficients usually exceed 0.90, suggesting almost perfect stability in rank order.

In short, the research on stability of individual differences in personality traits suggests that trait consistency across time is uniformly high. More specifically, results from a number of sources indicate that personality traits from all five factors are stable, that stability coefficients are higher for adults over age 30 than for younger adults, and that stability gradually decreases over long retest intervals. These findings fit neatly with those found for individual differences in intellectual functioning, as longitudinal studies on individual differences in intelligence also indicate that individual differences are stable over time, that stability coefficients are higher for adults over age 30 than for children or younger adults, and that stability gradually decreases over long retest intervals.

This page intentionally left blank.

5. Use of Tests in Selection

One of the aims of individual differences researchers is the prediction of future performance and/or behaviour. Psychological tests, attempting to measure individual differences related to occupational behaviours, have been used for over 60 years (Furnham, 1992). As aforementioned, Binet and his collaborators were the first to develop psychometric tests of intellectual functioning. Munsterberg, in 1913, brought this new testing movement into the industrial context in general, and employee selection in particular, by using a battery of tests available at the time to help select motormen for the Boston Railway Company (Furnham, 1992). Since then, psychological testing has blossomed fully and can be found in a number of settings such as industry, education, clinical, and counseling (Furnham, 1992).

The greatest boost for the use of tests for selection purposes came during World War I when 1.7 million American men were tested on either the Army Alpha or Army Beta for selection and classification purposes (Furnham, 1992). As stated by Dunnette and Borman (1979; cited in Kanfer, Ackerman, Murtha, and Goff, 1995), the use of these tests "marked the beginning of large-scale use of tests and other systematic methods to aid personnel decisions in the work of work" (p. 478). In addition to testing advances in the intelligence domain, the war brought about the development of the first standardized personality inventory: the Woodworth Test of Emotional Stability (Furnham, 1992). During the war, the Woodworth Test of Emotional Stability was administered to recruits for the purpose of screening out those who might be susceptible to wartime disorders and thus unsuitable for service (Furnham, 1992).

Following World War I, the use of tests spread to commercial organizations and by the middle of the 20th century, testing and selection procedures were considered so valid that psychologists were no longer required to administer them, instead, personnel managers were sufficient (Dillon & Watson, 1996).

Today, personality and mental ability tests are used as a selection device for any number of the following reasons: high number of applicants, high costs of interviewing, poor reliability of interviewing, necessity for objective, comparative data, or for allegations about favourtism. One of the benefits of psychological tests as a personnel selection strategy is that high selection validity translates into considerable financial savings for most organizations (Hunter & Hunter, 1984). For example, Hunter and Schmidt (1982; cited in Hunter & Hunter, 1984) estimated that, in 1980, productivity differences between complete use of cognitive ability tests and no use of the tests would amount to a minimum of $80 billion per year. That is, productivity differences due to the use or nonuse of ability tests would have been about as great as total corporate profits or about 20% of the total U.S. federal budget at that time.

A necessary condition for using traits for personnel selection and classification purposes is some degree of stability in the individual difference characteristics of the applicants (i.e., the predictors; Ilgen, 1994). Hence, scores on the predictors should be reliably stable over time. As shown in the previous section (i.e., section IV), individual differences in personality traits and individual differences in intellectual functioning remain very stable over time. As a result, individual differences in personality and intelligence serve as valid predictors of future behaviour. In the next section, evidence will be provided to indicate that individual

differences in personality and intelligence also serve as valid predictors of future job performance. (Note: the validity of the tests or assessment devices used in personnel selection is usually assessed in terms of the correlation between scores on a test and scores on some performance measure.)

5.1 Using Intelligence to Predict Work Performance

5.1.1 Civilian Studies

Hunter and Hunter (1984) reanalyzed Ghiselli's (1973) data on the relation between cognitive ability and performance and found that cognitive ability had a mean validity for performance ratings of 0.45 across all known job families. Although they found that there was no job for which cognitive ability does not predict performance ratings, they did find that the validity of general cognitive ability predicting performance dropped as complexity dropped. For example, they reported that the correlation between performance and cognitive ability decreased from 0.53 for managers, to 0.48 for service workers, to 0.46 for trades and crafts, to 0.28 for vehicle operators, and 0.27 for sales clerks. Thus, the more complex the job, the better cognitive ability predicts performance. It should be noted, however, that even the simplest occupation had a validity coefficient of 0.27, which is high enough to provide for substantial improvement in work force productivity if cognitive ability is used for selection (Hunter & Hunter, 1984).

Similar findings were reported by Hunter (1986), who meta-analyzed over 90 studies that used the General Aptitude Test Battery (GATB), developed by the U.S. Employment Service in 1945, as a predictor of performance ratings. Results also indicated that as complexity decreased, the validity of cognitive ability predicting performance ratings dropped from 0.58 for jobs classified as highly complex, to 0.51 for jobs of medium complexity, to 0.40 for jobs of low complexity. Furthermore, Hunter's meta-analysis indicated that psychomotor ability is relevant to many of the simpler jobs and has even higher validity than cognitive ability. He noted, however, that there were no jobs for which cognitive ability was not valid.

In addition to performance ratings, Hunter and Hunter's (1984) reanalysis of Ghiselli's (1973) work indicated that when general cognitive ability is used alone as a predictor, the average validity for training success across all jobs is 0.54. However, unlike performance ratings, there was no simple trend across levels of occupational complexity. For example, correlations of 0.71 (clerk), 0.66 (service worker), 0.65 (trades and craft), and 0.51 (managers) were obtained between cognitive ability and training success over a number of occupations.

Schmitt, Noe, Gooding, and Kirsch (1984) also found that cognitive ability was a valid predictor of training success ($r = 0.41$). Furthermore, ability was a better predictor of training success than biodata ($r = 0.25$), special aptitude ($r = 0.31$), and personality ($r = 0.17$).

Hunter and Hunter (1984) meta-analyzed a number of studies to examine the extent to which other variables were able to predict training success and proficiency. Their analysis indicated that cognitive ability predicted training and proficiency better than the other predictors. For example, in terms of training success, neither age (-0.01), education (0.20), college GPA (0.30), nor experience (0.01) were anywhere near as a valid predictor as cognitive ability.

Similar validity coefficients were found for the criterion of proficiency (age = -0.01, education = 0.10, college GPA = 0.11, experience = .18). In fact, job knowledge was the only predictor with a validity coefficient as high as cognitive ability ($r = 0.48$ for the criterion of performance). Hunter and Hunter (1984) noted, however, that the use of job knowledge tests are limited by the fact that they can only be used for prediction if the examinees are persons already trained for the job. As a result, they concluded that validity cannot be increased by replacing ability tests with any other known alternative predictor. They admit, however, that validity may be increased in some jobs by adding an appropriate second predictor.

5.1.2 Military Studies

The military services, like other employers, have used mental ability tests for selection and classification since the early part of this century. During WWI, American psychologists had already developed the Army Alpha and Beta tests to enable the screening of potential combatants (Gal, 1995). To date, the military is one of the largest consumers of personality assessment and intelligence tests (Gal, 1995). The main reason for this long history of use is that evidence shows that they work. For example, Kohs and Irle (1920; cited in Stogdill, 1974) conducted a follow-up study on the military careers of 116 college students and reported that intelligence was the best predictor of military success.

One of the most massive validation databases is the set of studies of training success conducted by the U.S. military (Hunter, 1986). Hunter (1986) summarized the data from over 500 studies for the four main job families (i.e., mechanical, clerical, electronic, and general technical) in military work that included 828 schools and nearly half a million military personnel and found that cognitive ability was a valid predictor for each of the job families. More specifically, correlations of 0.62, 0.58, 0.67, and 0.62 were found for the mechanical, clerical, electronic, and general technical job families, respectively. The only school for which validity was lower was combat, where the validity was about 0.45 as compared to the overall validity of 0.62. The validity coefficient for combat, however, is in almost perfect agreement with that found between training success and ability among civilians ($r = 0.54$). Thus, the military data show that cognitive ability predicts training success with high validity in every line of work.

In short, the available research indicates that mental ability predicts performance in a very wide range of human activities. Recent meta-analyses (Hunter & Hirsh, 1987; Hunter & Hunter, 1984) indicate that the average validity across all civilian jobs is 0.54 for a training success criterion and 0.45 for a job proficiency criterion when general cognitive ability alone is used as a predictor. Furthermore, this relationship

has been shown to be invariant across demographic groups such as sex, racial/ethnic, and age groups (Hawk, 1986). Validity is much higher, however, for complex jobs than for simple jobs. In addition, the military data show that cognitive ability predicts training success with high validity (0.62) in every line of work. Thus, the available evidence indicates that measures of ability can account for approximately 25% of the variance in performance for both civilian and military personnel.

5.2 Using Personality to Predict Work Performance

5.2.1 Civilian Studies

It has been traditionally believed that there is little evidence to support the use of personality measures in the personnel selection process. For example, Guion and Gottier (1965; cited in Goldstein & Lanyon, 1999) concluded their comprehensive review of the literature as follows: "it is difficult to advocate, with a clear conscience, the use of personality measures in most situations as a basis for making employment decisions" (p. 106).

Research in the last decade, however, has demonstrated that personality characteristics can be useful for predicting performance at work. Mount, Barrick, and Stewart (1998) attribute the support for personality traits as performance predictors to the recent emergence of the 'Big Five' theoretical scheme as the dominant model for conceptualizing personality. Recently, several reviews of the literature have concluded that personality measures, when classified with the 'Big Five' domains, are systematically related to a variety of criteria of job performance.

For example, a meta-analysis by Schmitt, Gooding, Noe, and Kirsch (1984) produced a sample-weighted mean correlation of 0.21 based on 32 personality scale validities using performance ratings as the criterion. They also found that although supervisory ratings were predicted less accurately than were status changes and wages, they were more predictable than productivity measures (e.g., sales volume).

Barrick and Mount's (1991) meta-analysis of 117 validity studies produced an overall corrected correlation of 0.11 between personality and job performance. In addition, they categorized the trait measures used in these studies into one of the 'Big Five' personality dimensions. As far as occupations are concerned, results indicated that Conscientiousness ($r = 0.22$) was a valid predictor across occupations and that the other personality factors only generalize their validity for some occupations. For example, Extraversion was a valid predictor for managers ($r = 0.18$) and sales ($r = 0.15$), Neuroticism was a valid predictor for police ($r = -0.10$), and Agreeableness was a valid predictor for police and managers ($r = 0.10$ for both). As far as criteria are concerned, training proficiency was predicted by Conscientiousness ($r = 0.23$) and Extraversion ($r = 0.26$), Openness to Experience ($r = 0.25$), and Agreeableness ($r = 0.10$), while job proficiency was predicted by Conscientiousness ($r = 0.23$) and Extroversion ($r = 0.10$).

Tett, Jackson, and Rothstein (1991) reported an overall corrected correlation of 0.24 between personality and job performance based on a meta-analysis of 86 studies conducted between 1968 and 1990. They also categorized the trait measures used in these studies into one of the 'Big Five' personality dimensions and reported obtained sample-weighted mean correlations of -0.22, 0.16, 0.27, 0.33, and 0.18 for Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. In addition, they found that the measurement interval was found to be unrelated to personality scale validity, indicating that the usefulness of personality measures in personnel selection remains stable over time. They also found that military job performance was more predictable than civilian job performance ($r = 0.30$ and 0.20, respectively).

Hough (1992; cited in Mount et al., 1992) used a nine-factor taxonomy to investigate relations with the criterion of teamwork. Three of the factors (Agreeableness, Emotional Stability [labeled adjustment], and Openness to Experience [labeled intellectance]) corresponded to the FFM level constructs and subtraits of the other two FFM dimensions were also examined (Extraversion was split into affiliation and potency, whereas Conscientiousness was split into achievement and dependability). In addition, the personality constructs of rugged individualism and locus of control were also examined.

Results indicated that Agreeableness ($r = 0.17$), the two Conscientiousness facets (achievement and dependability, $r = 0.14$ and $r = 0.17$), Emotional Stability ($r = 0.13$), and Openness to Experience ($r = 0.11$) were related to the criterion of teamwork. Thus, these results indicate at least three of the 'Big Five' personality factors, namely, Conscientiousness, Agreeableness, and Emotional Stability are moderate predictors of teamwork.

Similarly, Barrick, Stewart, Neubert, and Mount (1998) examined the relation between personality constructs and team outcomes at the team level, rather than the individual level. Results indicated that teams with higher mean levels of Conscientiousness ($r = 0.26$), Agreeableness ($r = 0.34$), and Emotional Stability ($r = 0.24$) were judged to be better performers than teams with lower mean scores for these personality factors.

Mount et al. (1998) conducted a meta-analysis on 11 studies that examined the validity of the FFM constructs for predicting performance in jobs that involve interpersonal interactions. Results indicated that Conscientiousness ($r = 0.26$), Agreeableness ($r = 0.21$), and Emotional Stability ($r = 0.18$), Openness to Experience ($r = 0.17$), and Extraversion ($r = 0.14$) were related to performance in jobs that involved interactions with others when the criterion was supervisor ratings of performance.

When the criterion was how well the employees interact with others, results indicated that Agreeableness ($r = 0.27$), Conscientiousness ($r = 0.20$), and Emotional Stability ($r = 0.19$), were the best FFM predictors. Hence, as with overall performance, Agreeableness, Conscientiousness, and Emotional Stability were found to be important predictors of performance ratings of interaction with others.

5.2.2 Military Studies

As found among civilian studies, studies of military personnel have also indicated that personality characteristics can be useful for predicting performance. For example, Tupes and Christal (1958; cited in Goldstein & Lanyon, 1999), using Air Force Academy cadets as their subjects, found clear and highly reliable evidence for five factors, which they labeled Surgency (Extroversion), Agreeableness, Conscientiousness, Emotional Stability, and Culture. More important, they found highly significant correlations between their five factor scores and a variety of performance measures obtained on these cadets, ranging from 0.60 for Conscientiousness to 0.24 for Extroversion.

In addition to their ability to predict performance ratings and training success, many researchers advocate the use of personality measures in making employment decisions because the use of well constructed personality measures do not appear to systematically discriminate against any ethnic or cultural group, and because the gender differences in norms do not seem to translate into differential selection rates for men and women applying for jobs (Hogan, Hogan, & Roberts, 1996).

Thus, the results of recent meta-analyses support a relation between the 'Big Five' personality scales and work performance. Tett et al. (1991) note, however, that any parsimony afforded by the use of broad frameworks, such as the 'Big Five', is purchased at the cost of diminished predictive power in relating particular personality traits to specific types of performance based on careful conceptual analyses and/or trait oriented job analysis. Thus, while they recognize the potential value of a macroscopic approach to personality validity, they conclude that personality scale validation is most appropriately undertaken through identification of specific trait-performance linkages.

5.3 Multidimensional Models

Sternberg (1994) has argued that no complete selection and assignment system can be based on only the measurement of general mental ability. He argued that one must also consider personality, interests, values, knowledge, motivation, roles and situations. Schmit (1994), however, regards this argument as unreasonable, as Sternberg is implying that one cannot study or use any predictor of performance unless one studies and uses all predictors of performance. While Schmit acknowledges that it would be nice to be able to measure every valid trait and take into account every situational contingency, he recognizes that it would be impossible to identify them all or get enough resources to research them all. Furthermore, there would not be enough testing time to measure them all. As a result, Schmit concludes that it is doubtful whether it will be possible any time soon to measure all relevant traits in any selection system.

To date, there is considerable evidence that cognitive ability appears to be relevant to predicting performance in virtually every job. In addition, evidence exists suggesting that broad personality traits are also relevant to predicting success in a wide array of jobs (for empirical evidence, see previous sections). As a result, is it possible that a combination of

cognitive ability and personality measures capture variance that is not adequately captured by even the best measures of ability or personality considered alone (Murphy & Shiarella, 1997).

Evidence supporting the use of multiple predictors stems from the work of Salgado (1998), who found that conscientiousness and emotional stability had incremental validity over general mental ability in the prediction of job performance. More specifically, he found that conscientiousness and emotional stability generalized validity across criteria and civil and military occupations, adding to total validity a percentage of about 10% to 11%.

McHenry, Hough, Toquam, Hanson, and Ashworth (1990) also demonstrated that some personality measures can provide substantial incremental validities over cognitive measures for the prediction of a variety of job related criteria. They found that the use of temperament-personality scales improved the prediction of criteria such as effort and leadership, personal discipline, physical fitness, and military bearing. They concluded that personality measures are "critical components of overall performance and should not be overlooked by a personnel selection and classification system" (McHenry et al., 1990, p.353).

Barrick, Mount, and Strauss (1993) examined the extent to which a multivariate model could predict job performance. In their study, a highly significant multiple correlation coefficient of 0.67 was found when mental ability and Conscientiousness, one of the 'Big Five' personality traits, was used to predict performance ratings of sales representatives of a large appliance manufacturer. Furthermore, the two predictors only correlated –0.07 with each other, suggesting that each tapped different aspects of future performance, which they labeled ability and motivation.

Hogan (1994; cited in Schmit, 1994) described considerable progress that has been made recently in the study of physical abilities. Hogan showed that the three major factors of physical ability (i.e., muscular strength, endurance, and movement quality) have substantial criterion related validity, particularly for job sample criterion measures of performance among military recruits. Unfortunately, however, few studies have examined the extent to which physical ability adds incremental validity above and beyond that already captured by mental ability. Hogan adds, however, that if only a minority of recruits are to be assigned to physically demanding jobs, the use of physical abilities tests for job classification (rather than initial selection) could have substantial utility. Hence, he concludes that the use of physical abilities measures in classification but not in selection avoids the possibility that potential recruits high on other desirable traits, cognitive ability in particular, would be screened out at selection because of mediocre physical ability.

In short, there is considerable evidence that cognitive ability appears to be relevant to predicting performance in virtually every job. In addition, evidence exists suggesting that broad personality traits are also relevant to predicting success in a wide array of jobs. As a result, a combination of cognitive ability and personality measures are able to capture variance that is not adequately captured by even the best measures of ability or personality considered alone. Hence, while it has been shown that cognitive ability is the best single predictor of performance, the addition of personality measures can provide substantial incremental validities over cognitive measures for the prediction of a variety of job related criteria.

This page intentionally left blank.

6. Individual Differences and Leadership

This section will deal with the relation between individual differences in personality, intelligence, and leadership. It should be acknowledged, however, that an adequate conceptualization of leadership must go far beyond the psychological realm (Gough, 1994). For example, Gough (1994) noted that there are cognitive, experiential, familial, morphological, physiological, and situational factors related to the occurrence of leadership.

In the early 1900's, leadership traits were studied to determine what made certain people great leaders. The theories developed were called "great man" theories because they focused on identifying the innate qualities and characteristics possessed by great social, political, and military leaders (Northouse, 1997). During this era, it was believed that great men were born not made.

Early in the 20th century, the "great man" theories evolved into trait theories. Trait theories did not make assumptions about whether leadership traits were inherited or acquired; instead, they concentrated on determining the specific traits that clearly differentiated leaders from followers (Northouse, 1997).

Trait theories were brought into question by leadership researchers as a result of the analyses of Mann (1959) and Stogdill (1948). Stogdill's review, which surveyed the literature on leadership from 1904 to 1947, found that no consistent set of traits differentiated leaders from nonleaders across a variety of situations. Similarly, Mann's review of the literature published between 1900 and 1957 on the relation between personality characteristics and leadership also concluded that there were low positive relationships between leadership and seven personality characteristics (introversion-extraversion, dominance, interpersonal sensitivity, masculinity-femininity, conservatism, intelligence, and adjustment).

Many recent researchers, however, have argued that the Mann and Stogdill reviews have been misinterpreted and that there are both theoretical and methodological reasons for reconsidering the relations between the traits and leaders. First, Lord et al. (1986) argued that the relations between the traits and leaders should be reconsidered because both reviews dealt with only leadership emergence or the perception of leadership in groups with no formal leader. Second, Lord et al. argued that there were many consistently significant positive relationships between personality and leadership emergence in both reviews that were seriously undermined. For example, even though Mann (1959; cited in Lord et al., 1986) reported that 88% of the 196 relations between intelligence and leadership were positive, 92 of the 196 relations were significant, and 99% of the significant relations were in the positive direction, he chose to emphasize the low median correlations between leadership and traits such as intelligence rather than the consistency in the trends he uncovered. Third, Lord et al. argued that methodologically, there have been a number of important developments in statistically aggregating results across studies since the Mann and Stogdill reviews were published and that techniques such as meta-analysis give better estimates of population parameters than the median correlation used in these earlier reviews.

6.1 Leadership and Intelligence: Empirical Findings

Northouse (1997) recently summarized a number of studies and concluded that intelligence is consistently reported by other researchers as being central to leadership. To date, most studies of leadership have found that intelligence is positively related to leadership. For example, Lord et al.'s (1986) meta-analysis on much of the previous trait research found a correlation of 0.52 between perceptions of leadership and intelligence. They concluded intelligence is a key characteristic in predicting leadership perceptions.

As aforementioned, Mann's review of the literature between the period of 1900 and 1957 (1959; cited in Lord et al., 1986) reported that 88% of the 196 relations between intelligence and leadership were positive, 92 of the 196 relations were significant, and 99% of the significant relations were in the positive direction.

Stogdill's (1948) results indicated that, in nearly all of the studies examined, the leader surpassed the average member of his or her group in intelligence. More specifically, among the 28 studies he summarized, 23 found that the leaders were brighter than the subordinates whereas only 5 found no significant differences.

One of the most significant findings concerning the relation of intelligence to leadership is the fact that extreme discrepancies between the intelligence of potential leaders and their followers work against the exercise of leadership (Stogdill, 1974). For example, Hollingworth (1929; cited in Stogdill, 1974) found that "among children with a mean IQ of 100, the IQ of the leader is likely to fall between 115 and 130 IQ. That is, the leader is likely to be more intelligent, but not too much more intelligent than the average of the group led." Hollingsworth's results also indicated that a child with an IQ of 160 has very little chance of being a popular leader in a group of children of average intelligence, but may become a leader in a group of children with a mean IQ of 130.

Stogdill (1974) explained this finding as being due to the leader's inability to communicate effectively with his or her subordinates. He argued that the average subordinate cannot comprehend a large part of the vocabulary employed by a leader of unusually superior intelligence, and that differences in interests, goals, and activity patterns also act as barriers to joint participation, which is a necessary condition for leadership. Thus, while intellectual ability is positively related to leadership, results indicate that a leader's intellectual ability should not vary too much from that of his or her subordinates.

In short, as found among the literature on personnel selection and classification, intelligence is the best single predictor of leadership performance.

6.2 Leadership and Personality: Empirical Findings

Gough (1994) examined the extent to which items on an adjective checklist correlated with leadership ratings among 95 male and 98 female college students. Results indicated that the three items with the largest positive correlations with leadership ratings were: assertiveness (r = 0.84 vs. 0.85), outspokenness (r = 0.80 vs. 0.76), and confidence (r = 0.74 vs. 0.72) for both males and females, respectively. The three items with the largest negative correlations with

leadership ratings were: shyness (r = -0.65 vs. –0.81), silence (r = -0.67 vs. -0.72), and being timid (r = -0.63 vs. -0.68) for both males and females, respectively.

Using the same group of subjects, Gough (1994) examined the extent to which the scales of the California Psychological Inventory (CPI) predicted leadership ratings. While results indicated that the CPI scales of dominance (r = 0.38), self-acceptance (r = 0.34), and Independence (r = 0.34) were the best single predictors for the overall group, the extent to which the CPI scales predicted leadership ratings varied for males and females. For example, for females, the three scales with the largest correlations with leadership ratings were empathy (r = 0.47), self-acceptance (r = 0.46), and independence (r = 0.41), whereas for males, the three scales with the largest correlations with leadership ratings were dominance (r = 0.40), independence (r = 0.27), and self-acceptance (r = 0.24). Gough noted that Megargee and Carbonell (1988) also reported that dominance was the strongest indicator of leadership potential in the CPI.

Gough (1994) also examined the extent to which the items of the CPI correlated with leadership potential of over 5000 subjects consisting of high school students, college students, graduates from a U.S. military academy, and a group of adults taking part in a personality assessment project. Results indicated positive median correlations (across samples) between leadership potential and CPI scales of dominance (r = 0.33), independence (r = 0.28), self-acceptance (r = 0.26), and empathy (r = 0.25). Gough also conducted a regression analysis to determine the extent to which a combination of scales may exceed the level of prediction afforded by dominance alone. Results indicated that a combination of dominance, social presence, independence, and empathy increased the validity to 0.40.

Lord et al. (1986) conducted a meta-analysis on much of the previous trait research and found that personality traits were strongly associated with individuals' perceptions of leadership. More specifically, they found that perceptions of leadership were correlated with masculinity-femininity (r = 0.34), conservatism (r = 0.22), adjustment (r = 0.21), dominance (r = 0.17), and extroversion (r = 0.15). These authors argued strongly that personality traits could be used to make discriminations consistently across situations between leaders and nonleaders.

In short, several studies have shown that individual differences in personality traits can serve as valid predictors of leadership criterion measures. More specifically, the trait of dominance, as measured by the CPI, has been shown to be the single best predictor of leadership criterion measures. Furthermore, multiple regression analyses have shown that a combination of personality traits may exceed the level of prediction afforded by dominance alone; this increase in validity, however, is far from substantial.

6.3 Leadership and Gender

To date, evidence indicates that there are few differences between the actual behaviours of male and female leaders (Schein, 1995). While research does show that some differences exist between males and females, it also indicates that there are fewer differences than commonly believed, suggesting instead that the differences within each sex are greater than the differences between the sexes (Schein, 1995). For example, a major investigation by researchers at the Centre for Creative Leadership found that executive men and women seem

to be virtually identical psychologically, intellectually, and emotionally (Schein, 1995). Hence, it appears that the traits necessary for effective leadership are prevalent among both male and female leaders.

6.4 Criticisms of the Trait Approach

Northouse (1997) criticizes the trait approach for the following two reasons: the failure to produce a definitive list of leadership traits and the failure to look at traits in relation to leadership outcomes. His first criticism stems from the fact that although an enormous number of studies have been conducted over the past century, the findings from these studies have been ambiguous and often uncertain. As a result, the trait approach has resulted in a list of traits that appears almost endless.

Northouse's (1997) second criticism arises from the fact that researchers have focused on the link between specific traits and leader emergence and have not tried to link leader traits with other outcomes such as productivity or employee satisfaction. For example, he argues that trait research does not provide data on whether leaders who might have high intelligence and strong integrity have better results than leaders without these traits.

In short, during the past century, researchers have attempted to determine the specific traits that clearly differentiate leaders from followers. Results from many studies indicate that many traits contribute to leadership. Northouse's (1997) summary suggests that intelligence, self-confidence, determination, integrity, and sociability are the main traits that are consistently identified in most of the available studies. In addition, the work by Gough (1994) and Megargee and Carbonell (1988) suggests that dominance is also a strong indicator of leadership potential. As found in the literature on personnel selection, intellectual ability is the single best predictor of leadership potential.

7. Individual Differences and Creativity

7.1 Creativity and Intelligence

To date, the available evidence indicates a positive association between creativity and intelligence (Eysenck, 1995). Of greater interest, however, is the finding that intelligence is only correlated with creativity measures up to IQ values of approximately 120; when this value is reached or exceeded, no correlation is typically found (Eysenck, 1995). Eysenck interprets this to mean that a certain amount of intelligence is needed to lay a foundation in knowledge that enables one to understand the fundamentals of the problem, interpret the rules, and give solutions that are socially acceptable.

7.2 Creativity and Personality

In their summary of more than two dozen studies, Dellas and Gaier (1970; cited in Eysenck, 1995) found a common pattern of personality traits among creative persons. More specifically, they found that creative individuals were characterized by independence in attitude and social behaviour, dominance, introversion, openness to stimuli, wide interests, self-acceptance, intuitiveness, flexibility, social presence and praise.

Similarly, Welsh (1975; cited in Eysenck, 1995) found a list of personality characteristics that characterized creative persons. For example, creative individuals were found to be original, adventurous, liberal, tolerant, spontaneous, flexible, and artistic. In addition to these positive traits, however, a number of negative traits were also found to characterize creative individuals. Thus, creative persons were also found to be unstable, irresponsible, disorderly, rebellious, self-seeking, uncooperative, impulsive, and careless.

Of even greater interest, however, is the finding that there is a large body of evidence linking creativity and genius with psychopathology (Eysenck, 1995). For example, Karlsson (1970; cited in Eysenck, 1995), on the basis of biographical material, claimed to have found the rate of psychosis to be 30% for great novelists, 35% for great poets, 35% for great painters, 25% for great mathematicians, and 40% for great philosophers, values that are significantly greater than that found among ordinary people (approximately 2%). Similarly, Andreasen (1970; cited in Eysenck, 1995), in a controlled study of 30 eminent writers and 30 matched control subjects, found that no fewer than 80% of the writers had experienced an episode of affective disorder, as compared to only 30% of the controls.

Eysenck (1995) also noted that when highly creative subjects are give personality questionnaires, their answers tend to be similar to those of neurotic or psychotic individuals, although usually at a lower level. In addition, Eysenck cited several studies that support a relation between psychoticism and creativity, noting that psychosis should never be identified with psychoticism, as the former is an illness and the latter a predisposition.

Eysenck (1995) argued the relation between creativity and psychoticism stems from the fact that both share certain mental processes or cognitive styles. In particular, he argues that

creativity and psychoticism consist of overinclusive or allusive thinking. This type of thinking is the product of a flat associative gradient that allows remote associations to influence cognitive processes of problem solving. He argued that psychoticism differs from psychosis by not being pathological, hence enabling people to use remote associations in a constructive way, whereas psychotics are overwhelmed by overinclusive thoughts and cannot cope with them in a critical manner.

Evidence supporting this theory stems from the work of MacKinnon (1962, cited in Eysenck, 1995), who found that among a group of creative, somewhat creative, and noncreative architects, the most creative group scored the highest and the noncreative group scored the lowest on a word association test.

In short, findings on the relation between individual differences in personality, intelligence, and creativity are far from clear. Eysenck (1995) has criticized the available findings as being "largely descriptive, full of anecdotal evidence, and without close links with the two disciplines of scientific psychology—the experimental and the psychometric" (p.231). Hence, more empirical findings are needed to determine whether creativity is associated with individual differences variables such as intelligence and personality characteristics.

8. Summary and Conclusions

Differential psychology refers to the objective and quantitative investigation of individual differences in behaviour (Anastasi, 1958). While the notion that individual differences exist can be dated back to Plato in 400 B.C., the systematic study of individual differences and psychometrics did not begin until the early 1900's. The modern period of differential psychology may be viewed as beginning with the work of Galton, who contributed to the field of individual differences research by advocating the measurement and quantification of both physical and mental traits on a large scale. Galton's view that a measure of intellectual functioning could be obtained through tests of sensory discrimination and RT, however, was criticized by Binet and Henri (1895) as being too sensory and as concentrating on simple, specialized abilities. Shortly afterwards, Binet and Henri invented the first practically useful test of intelligence, which was the progenitor of all subsequent individual tests of intelligence and profoundly influenced the item contents of group tests as well.

The greatest boost for the use of tests for selection purposes came during World War I when 1.7 million American men were tested on either the Army Alpha or Army Beta for selection and classification purposes (Furnham, 1992). To date, mental tests provide a primary method for studying individual differences. They are used for placing children in schools, promotion, educational guidance, selection of students for college, and selection of workers for industry. In addition, intelligence tests may be regarded as the most important innovation in developing and legitimating the idea of individual differences, as intelligence tests were able to scientifically demonstrate the existence of such differences.

Psychometrics is the branch of psychology concerned with the measurement of individual differences. To assess individual differences, one needs tests that are highly reliable. Reliability is an important issue for individual differences researchers who depend on correlational analyses to support their theories, as the reliability coefficients of the tasks in question will limit the maximum correlation that one can expect to find between the two tasks.

Another statistical tool that is important to individual differences researchers is factor analysis. Factor analysis can be used to construct tests, to resolve theoretical disputes about the number and nature of factors measured by tests, and to check whether tests work as they should or whether different tests should be used for different populations or cultures (Cooper, 1998).

Studies that have examined the causes of individual differences have found that both genetic and environmental factors influence the development of ability and personality traits. It should be noted, however, that for some traits, individual differences appear to be primarily shaped by genetic factors. For example, Vernon's (1979) summary of the available evidence suggests that 60% of the variability in intellectual functioning can be attributed to genetic influences. Similarly, Zuckerman (1991) has found that the main personality traits of extraversion, neuroticism, and psychoticism are also primarily shaped by genetic factors. In addition, the finding that genetic components become more influential with age also attests to the overwhelming influence that genes exert on individual differences.

Results from a number of longitudinal studies attest to the long-term stability of individual differences in intellectual functioning. Although it has been shown that mean levels of intelligence test scores decline slightly with age, individuals do not change in terms of their rank ordering of test performance. Furthermore, the rank ordering of individual differences in intellectual functioning has been shown to remain stable for intervals of 45 years and more (Arbuckle et al., 1998).

Similarly, the research on stability of individual differences in personality traits suggests that trait consistency across time is uniformly high. More specifically, results from a number of sources indicate that the 'Big Five' personality factors of Costa and McCrae (1992) are stable, that stability coefficients are higher for adults over age 30 than for younger adults, and that stability gradually decreases over long retest intervals. These findings fit neatly with those found for individual differences in intellectual functioning, as longitudinal studies on individual differences in intelligence also indicate that individual differences are stable over time, that stability coefficients are higher for adults over age 30 than for children or younger adults, and that stability gradually decreases over long retest intervals.

Because individual differences in personality characteristics and intelligence show temporal stability, personality and intelligence measures can serve as valid predictors of future behaviour. Today, personality and mental ability tests are used as selection devices to screen out unacceptable applicants and to reduce the high costs of interviewing. Hence, one of the benefits of psychological tests as a personnel selection strategy is that high selection validity translates into considerable financial savings for most organizations. This is especially important for the military, where the cost of training is extremely high. As a result, the military is one of the largest consumers of personality assessment and intelligence tests (Gal, 1995).

To date, there is considerable evidence that intellectual ability is the single best predictor of performance in virtually every job. In addition, evidence suggests that broad personality traits are also relevant to predicting success in a wide array of jobs. As a result, it is possible that a combination of cognitive ability and personality measures can capture variance that is not adequately captured by even the best measures of ability or personality considered alone. More research, however, is needed in this area. For example, no studies were found that examined the extent to which the various predictors of job performance or training interacted with one another nor have any studies that have used multivariate models examined the extent to which the addition of a cubed or quadratic function of a single variable adds to the predictability of the model (Simonton, 1995).

Research on leadership has also found that intellectual ability is the single best predictor of leadership perceptions. Research has shown that although effective leaders are typically brighter than their subordinates, they should not be too much brighter. Stogdill (1974) explained this finding as being due to the leader's inability to communicate effectively with his or her subordinates.

Results from studies that have attempted to determine the specific traits that clearly differentiate leaders from followers have found that intelligence, self-confidence, determination, integrity, and sociability are the main traits that consistently identify effective leaders (Northouse, 1997). As found with the personnel selection literature, more research is

needed in the area of multivariate models and the prediction of leadership, as the extent to which the various predictors of leadership interact with one another and the extent to which the addition of a cubed or quadratic function of a single variable adds to the predictability of leadership effectiveness is not clear (Simonton, 1995). In addition, future research is needed to determine how leaders' traits affect the outcomes of groups and teams in organizational settings, and which leader traits predict outcomes such as productivity and employee satisfaction.

Finally, the extent to which individual differences in intelligence and personality characteristics predict creativity is far from clear. Eysenck (1995) has criticized the available findings as being "largely descriptive, full of anecdotal evidence, and without close links with the two disciplines of scientific psychology—the experimental and the psychometric" (p.231). To date, preliminary evidence suggests that creative individuals are characterized as being intelligent, independent, introverted, and flexible, as well as unstable, irresponsible, disorderly, and careless.

A large body of evidence exists supporting a link between creativity and psychopathology. Eysenck (1995) argues that the relation between creativity and psychoticism stems from the fact that both share certain mental processes or cognitive styles, such as overinclusive or allusive thinking. More empirical findings are needed, however, to shed light on the relation between creativity, intelligence, and personality characteristics.

One weakness of the studies that have used individual differences in intellectual ability or personality characteristics as predictors of various criterion measures (i.e., performance ratings in the workplace, leadership rating, etc.) is that the researchers have assumed a linear relation between the predictor and the criterion (i.e., r measures the linear relation between two variables). In fact, most of the studies cited in this paper have failed to examine the scatterplots (i.e., between the various predictors and criterion measures) for curvilinearity. Because the correlation coefficient can seriously underestimate the relationship between criterion and predictor whenever the relation departs from linearity, and because the bulk of the research has relied heavily on linear measures of statistical association, it follows that the empirical literature may seriously underestimate the predictive value of many measures of personal attributes (Simonton, 1995). For example, the role of intelligence in leadership illustrates this predicament very well as further examination of the scatterplots reveals that the relation between intelligence and leadership is not linear; instead, beyond a certain level of intellect, further increases in cognitive ability can actually inhibit leader effectiveness (Simonton, 1995).

Thus, it is important to recognize that curvilinear or nonmonotonic functions may exist between various predictor and criterion variables when examining the relation between job performance (or leader effectiveness) and personal attributes. The more complex the functions are, the more our simple correlation coefficient may underestimate the predictive efficiency of intellectual ability and personality variables (Simonton, 1995).

In sum, individual differences in intelligence and personality characteristics can be used to predict training success and performance ratings for both military and civilians. To date, individual differences in intellectual ability appears to be the single best predictor of success for both personnel selection and leadership ratings. More research is needed, however, to

examine the extent to which curvilinear or nonmonotonic functions may exist between various predictor and criterion variables. In addition, the use of multivariate models that include both ability and personality measures, as well as more complex models that include the addition of a cubed or quadratic function of a single variable, will likely improve our ability to predict job performance and future leaders.

Table 1. *Summary of Median Correlations Between IQ's of Different Kinship Pairings Collected by Erlenmeyer-Kimling and Jarvik (1963)*

KINSHIP RELATION	NUMBER OF STUDIES	MEDIAN CORRELATION	GENETIC EXPECTATION
Monozygotic twins reared together	14	0.87	1.00
Monozygotic twins reared apart	4	0.75	1.00
Dizygotic same-sex twins reared together	11	0.56	0.50
Dizygotic unlike-sex twins reared together	9	0.49	0.50
Full siblings reared together	36	0.55	0.50
Foster parent with adopted child	3	0.20	0.00
Unrelated children reared together	5	0.24	0.00
Unrelated children reared apart	4	-0.01	0.00

(adapted from Table 11.1 from Vernon, 1979).

References

Alwin, D.F. (1994). Aging, personality, and social change: The stability of individual differences over the adult life span. In D. Featherman, R. Learner, & M. Perlmutter (Eds.), *Life-Span Development and Behaviour, Volume 12* (pp. 136-177). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Anastasi, A. (1958). *Differential Psychology: Individual and Group Differences in Behaviour* (3rd ed.). New York: Macmillan.

Anastasi, A. (1985). The use of personality assessment in industry: Methodological and interpretive problems. In H.J. Bernardin & D.A. Bownas (Eds.), *Personality Assessment in Organizations* (pp. 1-20). New York: Praeger Publishers.

Andreasen, N.C. (1987). Creativity and mental illness: Prevalence rates in writers and their first-degree relatives. *American Journal of Psychiatry, 144*, 1288-1292.

Angoff, W.H. (1988). The nature-nurture debate: Aptitudes and group differences. *American Psychologist, 43(9)*, 713-720.

Arbuckle, T.Y., Maag, U., Pushkar, D., & Chaikelson, J.S. (1998). Individual differences in trajectory of intellectual development over 45 years of adulthood. *Psychology and Aging, 13(4)*, 663-675.

Barrick, M.R., & Mount, M.K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Barrick, M.R., Mount, M.K. , & Strauss, J.P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78(5)*, 715-722.

Barrick, M.R., Stewart, G.L., Neubert, M.J., & Mount, M.K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology, 83*, 377-391.

Bayley, N. (1949). Consistency and variability in the growth of intelligence from birth to eighteen years. *Journal of Genetic Psychology, 75*, 165-196.

Bacharach, V.R., & Baumeister, A.A. (1998). Direct and indirect effects of maternal intelligence, maternal age, income, and home environment on intelligence of preterm, low-birth-weight children. *Journal of Applied Developmental Psychology, 19(3)*, 361-375.

Binet, A., & Henri, V. (1895). La psychologie individuelle. *L'Année psychologique, 2*, 411-463.

Binet, A., & Simon, T. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectual des anormaux. L'Année psychologique, 11, 191-224.

Bornstein, M.H. (1989). Information processing (habituation) in infancy and stability in cognitive development. *Human Development, 32*, 129-136.

Bouchard, T.J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*, 1055-1059.

Bradway, K., & Thompson, C. (1962). Intelligence in adulthood: A twenty-five year follow-up. *Journal of Educational Psychology, 53*, 1-14.

Brunas-Wagstaff, J. (1998). *Personality: A Cognitive Approach.* New York, NY, USA: Routledge.

Buss, A., & Poley, W. (1979). *Individual Differences: Traits and Factors.* New York: Gardner Press Inc.

Cadoret, R., Cain, C., & Crowe, R. (1983). Evidence for gene-environment interaction in the development of adolescent antisocial behaviour. *Behaviour Genetics, 13(3)*, 301-310.

Campbell, J. (1956). *The Hero With a Thousand Faces.* Cleveland: World.

Cattell, J.M. (1947). *James McKeen Cattell – Man of Science, Volumes I & II.* Lancaster, Pa.: Science Press.

Conley, J.J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality, and self-opinion. *Personality and Individual Differences, 5(1)*, 11-25.

Conrad, H.S., & Jones, H.E. (1940). A second study of familial resemblance in intelligence: environmental and genetic implications of parent-child and sibling correlations in the total sample. *Yearbook of the National Society for Studies in Education, 39(II)*, 97-141.

Cooper, C. (1998). *Individual Differences.* London, England: Arnold.

Costa, P.T., & McCrae, R.R. (1998). Trait theories of personality. In D. Baron & M. Herson (Eds.), *Advanced Personality. The Plenum Series in Social Clinical Psychology* (pp. 103-121). New York: Plenum.

Costa, P.T., & McCrae, R.R. (1997). Longitudinal stability of adult personality. In R. Hogan, J. Johnson, & B. Briggs (Eds.), *Handbook of Personality Psychology* (pp. 269-290). San Diego, CA: Academic Press.

Costa, P.T., & McCrae, R.R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. Heatherton & J. Weinberger (Eds.), *Can Personality Change?* (pp. 21-40). Washington D.C.: American Psychological Association.

Costa, P.T., & McCrae, R.R. (1992). Trait psychology comes of age. In T.B. Sonderegger (Ed.), *Nebraska Symposium on Motivation, Volume 39* (pp. 169-204). Lincoln: University of Nebraska Press.

Costa, P.T., & McCrae, R.R. (1988). Personality in adulthood: A six year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54(5)*, 853-863.

Costa, P.T., & McCrae, R.R. (1977). Age differences in personality structure revisited: Studies in validity, stability, and change. *International Journal of Aging and Human Development, 8*, 261-275.

Costa, P.T., McCrae, R.R., & Arenberg, D. (1980). Enduring dispositions in adult males. *Journal of Personality and Social Psychology, 38*, 793-800.

Costa, P.T., Yang, J., & McCrae, R.R. (1998). Aging and personality traits: Generalizations and Clinical Implications. In I. Nordhus, G. Vandenbos, S. Berg, & P. Fromholt (Eds.), *Clinical Geropsychology* (pp. 33-48). Washington D.C.: American Psychological Association.

Cronbach, L.J. (1970). *Essentials of Psychological Testing, 3rd ed.* New York: Harper & Row.

Crow, J. F., Neel, J.V., & Stern, C. (1967). Racial studies: Academy states position on call for new research. *American Psychologist, 30*, 1-14.

Dellas, M., & Gaier, E.L. (1970). Identification of creativity: The individual. *Psychological Bulletin, 73*, 55-73.

Dillon, A., & Watson, C. (1996). User analysis in HCI: The historical lesson from individual differences research. *International Journal of Human-Computer Studies, 45(6)*, 619-637.

Dobreva-Martinova, T. (1999). Psychometric analysis of the unit climate profile questionnaire based on surveys of deployed Canadian Forces personnel (Directorate Human Resources Research and Evaluation Sponsor Research Report 99-3). Ottawa, Ontario: Department of National Defence Headquarters.

Dunnette, M.D., & Borman, W.C. (1979). Personnel selection and classification systems. *Annual Review of Psychology, 30*, 477-525.

Eysenck, H.J. (1995). Creativity as a product of intelligence and personality. In D.H. Saklofske & M. Zeidner (Eds), *International Handbook of Personality and Intelligence* (pp.231-247). New York, N.Y.: Plenum Publishing Corporation.

Feldman, D., & Adams, M. (1989). Intelligence, stability, and continuity: Changing conceptions. In M. Bornstein & N. Krasnegor (Eds.), *Stability and Continuity in Mental Development* (pp. 298-309). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Finn, S.E. (1986). Stability of personality self-ratings over 30 years: Evidence for an age/cohort interaction. *Journal of Personality and Social Psychology, 50*, 813-818.

Freeman, F.N., Holzinger, K.J., & Mitchell, B.C. (1928). The influence of environment on the intelligence, school achievement, and conduct of foster children. *Twenty-seventh Yearbook of the National Society for the Study of Education* (part 1; pp. 103-217).

Furnham, A. (1992). *Personality at Work: The Role of Individual Differences in the Workplace.* New York: Routledge.

Gal, R. (1995). Personality and intelligence in the military. *International Handbook of Personality and Intelligence* (pp.727-737). New York, N.Y.; Plenum Publishing Corporation.

Galton, F. (1869). *Hereditary Genius: An Inquiry Into Its Laws And Consequences.* London: Macmillan.

Galton, F. (1884). *Hereditary Genius.* New York: D. Appleton.

Ghiselli, E.E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology. 26*, 461-477.

Ghiselli, E.E., & Barthol, R.P. (1953). The validity of personality inventories in the selection of employees. *Journal of Applied Psychology, 37*, 18-20.

Gilliland, A.R., & Clark, E.L. (1939). *Psychology of Individual Differences.* New York: Prentice-Hall.

Gojone, H., & Stevenson, J. (1991). A longitudinal twin study of temperament and behaviour problems: Common genetic or environmental influences? *Journal of the American Academy of Child and Adolescent Psychiatry, 36(10)*, 1448-1456.

Goldstein, L.D., & Lanyon, R.I. (1999). Applications of personality assessment to the workplace: A review. *Journal of Business and Psychology, 13*, 291-322.

Gough, H.G. (1994). Testing for leadership with the California Psychological Inventory. In K. Clark & M. Clark (Eds.), *Measures of Leadership* (pp. 355-379). West Orange, New Jersey: Leadership Library of America.

Guerrin, D.W., & Gottfried, A.W. (1994). Developmental stability and change in parent reports of temperament: A ten-year longitudinal investigation from infancy through preadolescence. *Merrill-Palmer Quarterly, 40*, 334-355.

Guion, R.M., & Gottier, R.F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135-164.

Hawk, J. (1986). Real world implications of g. *Journal of Vocational Behavior, 29*, 411-414.

Hertzog, C., & Schaie, K.W. (1983). Stability and change in adult intelligence I. Analysis of longitudinal covariance structures. *Psychology and Aging, 1(2)*, 159-171.

Hofstaetter, P. (1954). The changing composition of "Intelligence": A study in T technique. *Journal of Genetic Psychology, 85*, 159-164.

Hogan, R., Hogan, J., & Roberts, B.W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist, 51*, 469-577.

Hollingworth, L.S. (1929). *Gifted Children.* New York: Macmillan.

Hough, L.M. (1992). The Big Five personality variables-construct confusion: description versus prediction. *Human Performance, 5*, 135-155.

Howell, D.C. (1992). *Statistical methods for psychology (3rd ed).* Boston, MA: Pws-Kent Publishing Co.

Hunter, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.

Hunter, J.E., & Hirsh, H.R. (1987). Applications of meta-analysis. In C.L. Cooper & I.T. Robertson (Eds), *International review of industrial and organizational psychology 1987* (pp. 321-357). Chichester, England UK: John Wiley & Sons.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.

Hunter, J.E., & Schmidt, F.L. (1982). Fitting people to jobs: The impact of personnel selection on national productivity. In M.D. Dunnette & E.A. Fleishman (Eds.), *Human Performance and Productivity: Human Capability Assessment* (pp. 233-284). Hillsdale, NJ: Erlbaum.

Ilgen, D.R. (1994). Jobs and roles: Accepting and coping with the changing structure of organizations. In M.G. Rumsey & C.L. Walker (Eds), *Personnel selection and classification* (pp. 13-32). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Jarvick, L.F., & Erlenmeyer-Kimling, L. (1967). Survey of familial correlations in measured intellectual functions. In J. Zubin & G. Jervis (Eds.), *Psychopathology of Mental Development.* New York: Grune and Stratton.

Jensen, A. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1-123.

Kanfer, R., Ackerman, P.L., Murtha, T., & Goff, M. (1995). In D. Saklofske & M. Zeidner (Eds*), International handbook of personality and intelligence. Perspectives on individual differences* (pp. 577-602). New York, NY, USA: Plenum Press.

Karlsson, J. (1970). Genetic association of giftedness and creativity with schizophrenia. *Heredity, 66*, 171-182.

Kohs, S.C., & Irle, K.W. (1920). Prophesying army promotion. *Journal of Applied Psychology, 4*, 73-87.

Loehlin, J. (1992). *Genes and Environment in Personality Development.* Newbury Park, CA: Sage Publications, Inc.

Loehlin, J.C., Lindzey, G., & Spuhler, J.N. (1975). *Race Differences in Intelligence.* San Francisco: Freeman.

Leon, G.R., Gillum, B., Gillum, R., & Gouze, M. (1979). Personality stability and change over a 30-year period: Middle age to old age. *Journal of Consulting and Clinical Psychology, 47*, 517-524.

Lord, R.G., DeVader, C.L., & Alliger, G.M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology, 71*, 42-410.

MacKinnon, D.W. (1962). The nature and nurture of creative talent. *American Psychologist, 17*, 484-495.

Mann, R.D. (1959). A review of the relationships between personality and performance in small groups. *Psychological Bulletin*, 56, 241-270.

Marks, R. (1976-1977). Providing for individual differences: A history of the intelligence testing movement in North America. *Interchange, 7(3)*, 3-16.

McCrae, R.R., & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175-215.

McGuffin, P., & Thapar, A. (1992). The genetics of personality disorder. *British Journal of Psychiatry, 160*, 12-23.

McHenry, J., Hough, L., Toquam, J., Hanson, M., & Ashworth, M. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology. 43*, 335-354.

Megargee, E.I., & Carbonell, J.L. (1988). Evaluating leadership with the CPI. In C.D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment, Volume 7* (pp. 203-219). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Miles, D., & Carey, G. (1997). Genetic and environmental architecture on human aggression. *Journal of Personality and Social Psychology, 72(1)*, 207-217.

Morton, N.E. (1972). Human behavioural genetics. In L. Ehrman, G.S. Omenn, & E. Caspari (Eds.), *Genetics, Environment, and Behaviour* (pp. 247-265). New York: Academic Press.

Mount, M.K., Barrick, M.R., & Stewart, G.L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165.

Munsinger, H. (1977). A reply to Kamin. *Behaviour Genetics, 7*, 407-409.

Murphy, G. (1938). *An Historical Introduction to Modern Psychology*. New York: Harcourt, Brace, & Company, Inc.

Murphy, K.R., & Shiarella, A. (1997). Implications of the multidimensional nature of job performance for the validity of selection testing: multivariate frameworks for studying test validity. *Personnel Psychology, 50*, 823-854.

Northouse, P. (1997). *Leadership: Theory and Practice.* Thousand Oaks, California: Sage Publications.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric Theory, 3rd Edition.* New York: McGraw Hill.

Owens, W. A. (1966). Age and mental abilities: A second adult follow-up. *Journal of Educational Psychology, 57*, 311-325.

Plomin, R. (1988). The nature and nurture of cognitive abilities. In R. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence, Volume 4* (pp. 1-33). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Roberts, B.W., & DelVecchio, W.F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126(1)*, 3-25.

Rosenzweig, M., Leiman, A., & Breedlove, S.M. (1996). Biological Psychology. Sunderland, Massachusetts: Sinauer Associates.

Salgado, J. (1998). Big five personality dimensions and job performance in army and civil occupations: A European Perspective. *Human Performance, 11*, 271-288.

Scarr-Salapatek, S. (1971). Unknowns in the IQ equation. *Science, 174*, 1223-1228.

Schaie, K.W. (1983). The Seattle longitudinal study: A 21-year exploration of psychometric intelligence in adulthood. In K.W. Schaie (Ed.), *Longitudinal Studies of Adult Development* (pp. 64-135). New York: Guilford.

Schaie, K.W., & Hertzog. (1983). Fourteen-year cohort-sequential analyses of adult intellectual development. *Developmental Psychology, 19(4)*, 531-543.

Schaie, K.W., & Strother. (1968). A cross-sequential study of age changes in cognitive behavior. *Psychological Bulletin, 70(6)*, 671-680.

Schein, V. (1995). Would women lead differently? In T. Wren (Ed.), *The Leaders Companion: Insights on Leadership Through the Ages*. New York, N.Y.: Free Press.

Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982. *Personnel Psychology, 37*, 407-421.

Schuerger, J.M., Zarrella, K.L., & Hotz, A.S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56(5)*, 777-783.

Schwartzman, A.E., Gold, D., Andres, D., Arbuckle, T.Y., & Chaikelson, J. (1987). Stability of intelligence: A 40-year follow up. *Canadian Journal of Psychology, 41(2)*, 244-256.

Simon, A., & Thomas, A. (1983). Means, standard deviations and stability coefficients on the EPI for Further Education and College of Education students. *Personality and Individual Differences, 4*, 95-96.

Simonton, D.K. (1995). Personality and intellectual predictors of leadership. *International Handbook of Personality and Intelligence* (pp.739-757). New York, N.Y.: Plenum Publishing Corporation.

Skodak, M., & Skeels, H.M. (1945). A follow-up study of children in adoptive homes. *Journal of Genetic Psychology, 66*, 21-58.

Skodak, M., & Skeels, H.M. (1949). A final follow-up study of one hundred adopted children. *Journal of Genetic Psychology, 75*, 85-125.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15(2)*, 201-293.

Stogdill, R. (1948). Personal factors associated with leadership. *Journal of Psychology, 25*, 35-71.

Stogdill, R. (1974). *Handbook of Leadership.* New York: The Free Press.

Terman, L.M. (1916). *The Measurement of Intelligence.* Boston: Houghton Mifflin.

Tett, R.P., Jackson, D.N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 703-742.

Thurstone, L.L. (1928). An experimental study of nationality preferences. *Journal of General Psychology, 1*, 405-425.

Tryon, R.C. (1940). Genetic differences in maze-learning ability in rats. *Yearbook of the National Society for the Study of Education, 39(I)*, 111-119.

Tupes, E.C., & Christal, R.E. (1958). Stability of personality trait rating factors obtained under diverse conditions. *USAF WADC Technical Note*, No. 58-61.

Tyler, L. (1995). The challenge of diversity. In D. Lubinski & R. Dawis (Eds.), *Assessing Individual Differences in Human Behaviour: New Concepts, Methods, and Findings* (pp. 1-13). Palo Alto, CA: Davies-Black Publishing.

Vernon, P. (1979). *Intelligence: Heredity and Environment.* San Francisco: W.H. Freeman and Co.

Watson, J.J.B. (1930). Behaviourism. New York: Norton.

**Welsh,

Williams, R. (1961). *The Long Revolution.* London: Chatto & Windas.

Woodruff, D.S. (1983). The role of memory in personality continuity: A 25 year follow-up. *Experimental Aging Research, 9*, 31-34.

Yerkes, R.M. (1921). Psychological examining in the United States army. In, *National Academy of Sciences (Volume 10).* Washington, D.C.: U.S. Government Printing Office.

Zuckerman, M. (1991). *Psychobiology of Personality.* New York: Cambridge University Press.

DOCUMENT CONTROL DATA SHEET

1a. PERFORMING AGENCY	2. SECURITY CLASSIFICATION
DCIEM	UNCLASSIFIED -
1b. PUBLISHING AGENCY	
DCIEM	

3. TITLE

(U) Individual Differences: A Review

4. AUTHORS

Stokes, Tonya L.

5. DATE OF PUBLICATION	6. NO. OF PAGES
March 1 , 2002	70

7. DESCRIPTIVE NOTES

8. SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY
Sponsoring Agency:
Monitoring Agency:
Contracting Agency :
Tasking Agency:

9. ORIGINATORS DOCUMENT NO.	10. CONTRACT GRANT AND/OR PROJECT NO.	11. OTHER DOCUMENT NOS.
Technical Report TR 2001-010		

12. DOCUMENT RELEASABILITY

Unlimited distribution

13. DOCUMENT ANNOUNCEMENT

Unlimited announcement

14. ABSTRACT

(U) Differential psychology refers to the objective and quantitative investigation of individual differences in behaviour. While the notion that individual differences exist can be dated back to Plato in 400 B.C., the systematic study of individual differences and psychometrics did not begin until the early 1900's. The following paper will: a) trace the history of individual differences research, b) explore the methods of measuring individual differences, c) identify the causes of individual differences, d) provide evidence attesting to the long term stability of individual differences, e) provide evidence that individual differences in intelligence and personality traits can be used to predict success in both civilian and military jobs, and f) provide evidence that individual differences in intelligence and personality traits can be used to predict leadership and creativity. Finally, this paper will highlight some of the criticisms of the current individual differences research and suggest future directions for researchers in this field.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) individual differences; measurement of individual differences; causes of individual differences; heredity versus environment; stability of individual differences in intelligence and personality; use of tests in selection; leadership & intelligence; leadership & personality; creativity & intelligence; creativity & personality

Defence R&D Canada
is the national authority for providing
Science and Technology (S&T) leadership
in the advancement and maintenance
of Canada's defence capabilities.

R et D pour la défense Canada
est responsable, au niveau national, pour
les sciences et la technologie (S et T)
au service de l'avancement et du maintien des
capacités de défense du Canada.

www.drdc-rddc.dnd.ca